Stochastic Processes and Financial Mathematics
(part two)

Dr Nic Freeman

January 28, 2022

Contents

Chapter 9

The transition to continuous time

Up until now, we have always indexed time by the integers, $n = 1, 2, 3, \ldots,$. For the remainder of the course, we will move into continuous time, meaning that our time will be indexed as $t \in [0, \infty)$. We need to update some of our terminology to match.

As before, we work over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 9.0.1 A stochastic process (in continuous time) is a family of random variables $(X_t)_{t=0}^{\infty}$. We think of $t \in [0, \infty)$ as time.

As in discrete time, we will usual write $(X_t) = (X_t)_{t=0}^{\infty}$. We will also sometimes write simply X instead of (X_t).

Definition 9.0.2 We say that a stochastic process (X_t) is *continuous* (or, equivalently, has *continuous paths*) if, for almost all $\omega \in \Omega$, the function $t \mapsto X_t(\omega)$ is continuous.

In words, we should think of a continuous stochastic process as a random continuous function. For example, if A, B and C are i.i.d. $N(0, 1)$ random variables, $X_t = At^2 + Bt + C$ for $t \in [0, \infty)$ is a random continuous (quadratic) function. In this course, we will usually be more interested in situations where, in some sense, randomness appears and causes the stochastic process to change value as time passes. To do so, we need to think about filtrations.

Definition 9.0.3 We say that a family $(\mathcal{F}_t)$ of σ-fields is a (continuous time) filtration if $\mathcal{F}_u \subseteq \mathcal{F}_t$ whenever $u \leq t$.

A stochastic process (X_t) is adapted to the filtration $(\mathcal{F}_t)$ if $M_t \in m\mathcal{F}_t$ for all $t \geq 0$.

In continuous time, our standard setup is that we will work over a *filtered space* $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $(\mathcal{F}_t)$ is a filtration. If we are given a stochastic process (X_t), implicitly over some probability space, the *generated (or natural) filtration* of the stochastic process is $\mathcal{F}_t = \sigma(X_u)\,;\, u \leq t)$.

Lastly, we upgrade our definition of a martingale into continuous time.

Definition 9.0.4 A (continuous time) stochastic process (M_t) is a martingale if

1. (M_t) is adapted,
2. $M_t \in L^1$ for all t,
3. $\mathbb{E}[M_t | \mathcal{F}_u] = M_u$ for all $0 \leq u \leq t$.

We say that (M_t) is a submartingale if, instead of 3, we have $\mathbb{E}[M_t|\mathcal{F}_u] \geq M_u$ almost surely. We say that (M_t) is a supermartingale if, instead of 3, we have $\mathbb{E}[M_t|\mathcal{F}_u] \leq M_u$ almost surely.

There are continuous time equivalents of the results (e.g. the optional stopping theorem) that we proved for discrete time martingales, but they are outside of the scope of this course.

Chapter 10

Brownian motion

In this chapter we study the most important example of a stochastic process: Brownian motion. In essence, Brownian motion is the continuous time equivalent of the symmetric random walk that we studied in Section 4.1.

10.1 The limit of random walks

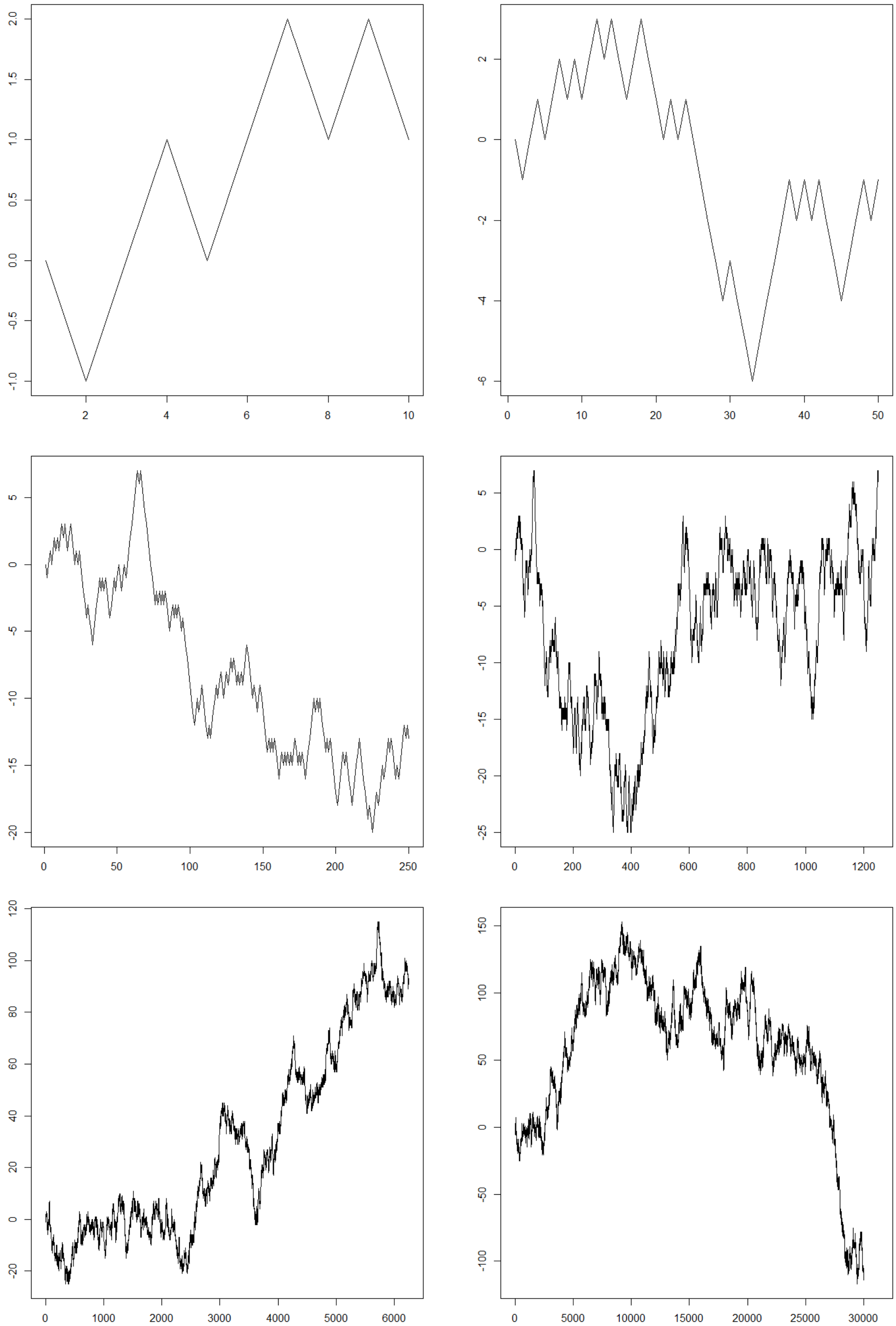
The discovery of Brownian motion has a distinguished place in the history of both science and mathematics. It is named after the botanist Robert Brown who, in 1827, through a microscope, saw erratic movements being made by tiny pollen organelles floating on water. The cause of these movements was explained later by Albert Einstein and the physicist Jean Perrin: the movements were caused by (the cumulative effect of) many individual water molecules hitting the tiny organelles. This realization provided the 'modern science' of the time with a key piece of evidence for the existence of atoms[1]. Around the same time, the american mathematician Norbert Wiener, building on earlier work of Louis Bachelier, developed a mathematical model for stock prices and independently discovered Brownian motion.

Today, Brownian motion is at the heart of many important models of the physical world. We will see some examples in future sections of the course; for now our first task is to construct the process.

Brown, Einstein and Perrin studied pollen movements on the surface of still water, meaning they observed movements in two (spatial) dimensions $\mathbb{R}^2$. Bachelier, by contrast, saw stocks prices moving up and down - in one dimension $\mathbb{R}$. In both cases the underlying principle is one of 'completely random' movement. We will restrict to the one dimensional case in this course.

Recall the symmetric random walk from Section 4.1. We will look at six pictures of (samples of) it, where in each picture the random walk has run for a successively longer time ($T = 10, 50, 250, 1250, 6250, 31250$). We fit each such picture into the same size box – we can think of this as zooming out, so in each inch of space on the paper we see more and more, smaller and smaller, jumps of the random walk. The results are intriguing:

[1] At the time, scientists were not confident of the existence of atoms; there were other competing theories that had not yet been disproved.



As we zoom out further and further, the pictures are starting to look very similar in character. The last three are all 'jagged' in a conspicuously similar way. If you look carefully, you can see that each picture contains precisely the first fifth (in terms of time passed) of the next picture. From the axis on pictures, we might guess that, if we have T units of time on the x-axis, we need about $\sqrt{T}$ units of space on the y-axis. This is not surprising: a short calculation (which we omit) shows that $\mathbb{E}[|X_T|] \approx \sqrt{T}$ as $T \to \infty$.

The fact that our pictures start to look very similar in character, as T gets larger, is highly suggestive: it suggests that as we keep scaling out we will see convergence to a *limit*. The limit, like the random walk, will be random; it will be a continuous time stochastic process.

In Chapter 6 we studied limits for random variables. This theory can be extended, into looking at limits of whole stochastic processes, because a stochastic process $(Z_t)_{t=0}^{\infty}$ is just the set of random variables $\{Z_t\,;\, t \in [0,\infty)\}$. We won't study modes of convergence for stochastic processes in this course, but hopefully the idea is clear. Various tools from analysis, that are outside the scope of our course, can be used to prove that a limit exists in this case – the limit is called *Brownian motion*, and it is the focus of this chapter.

Remark 10.1.1 You can reproduce the pictures yourself, using R, with the code for e.g. the first one:

```
> T=10
> set.seed(1)
> x=c(0,2*rbinom(T-1,1,0.5)-1)
> y=cumsum(x)
> par(mar=rep(2,4))
> plot(y,type="l")
```

10.2 Brownian motion

To work with Brownian motion mathematically, we need more than the pictures from the previous section. What we need is a theorem that (1) tells us that Brownian motion exists and (2) gives us some properties to work with. We begin our mathematical treatment of Brownian motion as follows, with a definition that it also an existence theorem:

Theorem 10.2.1 *There is a stochastic process (B_t) such that:*

1. *The paths of (B_t) are continuous.*
2. *For any $0 \le u \le t$, the random variable $B_t - B_u$ is independent of $\sigma(B_v\,;\, v \le u)$.*
3. *For any $0 \le u \le t$, the random variable $B_t - B_u$ has distribution $N(0, t-u)$.*

Further, any stochastic process which satisfies these three conditions has the same distribution as (B_t).

Definition 10.2.2 If $B_0 = 0$ we say that (B_t) is a *standard* Brownian motion.

From now on we fix some notation, which we will use for the remainder of the course:

> **We write (B_t) for a standard Brownian motion, and $\mathcal{F}_t = \sigma(B_u\,;\, u \le t)$ for its generated filtration. We work over the filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$.**

Let us record a few simple facts about Brownian motion that we will use repeatedly in later chapters. Putting $u = 0$ into the second property and noting that $B_0 = 0$, we obtain that the distribution of Brownian motion at time t is $B_t \sim N(0, t)$. Hence, also,

$$\mathbb{E}[B_t] = 0$$

for all t. In exercise 10.4 you are are asked to show that if $Z \sim N(0, t)$ then $\mathbb{E}[Z^2] = t$. Hence

$$\mathbb{E}[B_t^2] = \mathrm{var}(B_t) = t \tag{10.1}$$

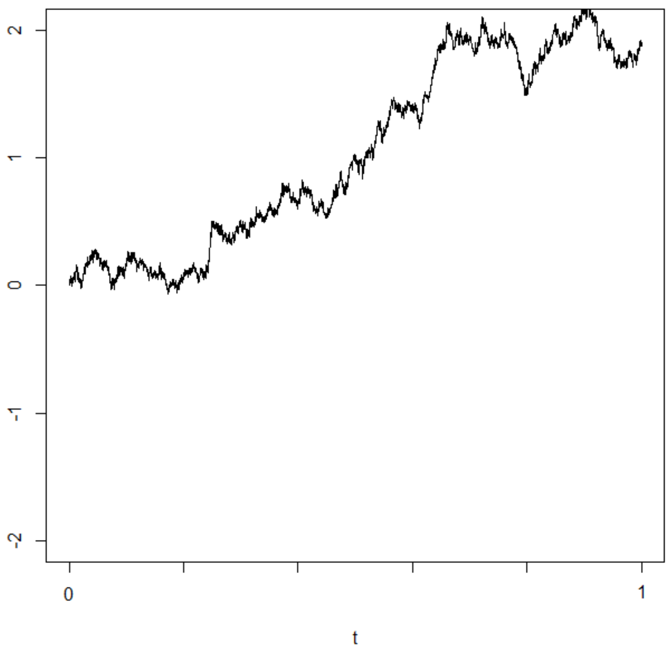
for all t. It is also useful to know that $B_t^n \in L^1$ for all $n \in \mathbb{N}$. See exercise 10.4 for a proof of this fact.

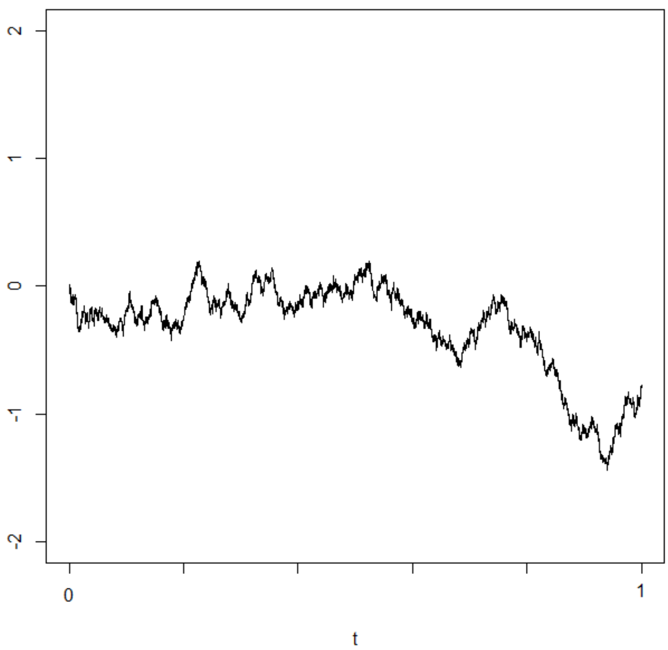
Along the same lines, another formula that we will use repeatedly (and that you should remember) is that if $Z \sim N(\mu, \sigma^2)$ then

$$\mathbb{E}[e^Z] = e^{\mu + \frac{1}{2}\sigma^2}. \tag{10.2}$$

This formula turns out to be surprisingly useful in situations involving Brownian motion.

Here are a couple of samples of (standard) Brownian motion. They look very similar in character to the pictures from Section 10.1, when T was large. The key point is that, now, instead of large T, time and space have been rescaled so as we only watch 1 unit of time.





You might like to note the similarity of these pictures to the jagged nature of Figure 1.1 (which was a plot of the stock price of Lloyds Banking Group), which we reproduce here for convenience:

In fact, we won't use Brownian motion for our stock price model; we'll use a slight modification known as 'geometric' Brownian motion. For now, we need to collect together some more information about Brownian motion, and develop our modelling tools further, but we'll return to the question of stock price models in Section 12.2 and Chapter 14.

10.3 Brownian motion and the heat equation

Brownian motion lies at the heart of many modern models of the physical world. Before we study Brownian motion in its own right, let us give one example of a model with close connection to Brownian motion, namely heat diffusion.

Consider a long thin metal rod. If, initially, some parts of the rod are hot and some are cold, then as time passes heat will diffuse through the rod: the differences in temperature slowly average out. Suppose that the temperature of the metal in the rod at position x at time t is given by $u(t,x)$, where $t \in [0,\infty)$ represents time and $x \in \mathbb{R}$ represents space. Suppose that, at time 0, the temperature at the point x is $f(x) = u(0,x)$.

Then (as you may have seen from e.g. MAS222), it is well known that the *heat equation*

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} \tag{10.3}$$

with the initial condition

$$u(0,x) = f(x) \tag{10.4}$$

describes how the temperature $u(t,x)$ changes with time, from an initial temperature of $f(x)$ at site x. This equation has a close connection to Brownian motion, which we now explore.

Remark 10.3.1 ($\star$) In the world of PDEs, the factor $\frac{1}{2}$ in (10.3) is not normally included, but in probability we tend to include it. The difference is just that time runs twice as fast (i.e. we substituted $2t$ in place of t), which isn't very important.

If we start Brownian motion from $x \in \mathbb{R}$, then $B_t \sim x + N(0,t) \sim N(x,t)$, so we can write down its probability density function

$$\phi_{t,x}(y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right).$$

Lemma 10.3.2 *$\phi_{t,x}(y)$ satisfies the heat equation* (10.3).

Proof: Note that if any function $u(t,x)$ satisfies the heat equation, so does $u(t,x-y)$ for any value of y. So, we can assume $y=0$ and need to show that

$$\phi_{t,x}(0) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right)$$

satisfies (10.3). This is an exercise in partial differentiation. Using the chain and product rules:

$$\begin{aligned}
\frac{\partial \phi}{\partial t} &= \frac{\frac{-1}{2}}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right) + \frac{1}{\sqrt{2\pi t}}\frac{-x^2(-1)}{2t^2} \exp\left(-\frac{x^2}{2t}\right) \\
&= \frac{1}{\sqrt{2\pi}}\left(\frac{x^2}{2t^{5/2}} - \frac{1}{2t^{3/2}}\right) \exp\left(-\frac{x^2}{2t}\right)
\end{aligned}$$

and

$$\frac{\partial \phi}{\partial x} = \frac{1}{\sqrt{2\pi t}}\frac{-2x}{2t} \exp\left(-\frac{x^2}{2t}\right)$$

$$= \frac{-x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right)$$

so that

$$\begin{aligned}\frac{\partial^2 \phi}{\partial x^2} &= \frac{-1}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right) + \frac{-x}{\sqrt{2\pi t^3}} \frac{-2x}{2t} \exp\left(-\frac{x^2}{2t}\right) \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{x^2}{t^{5/2}} - \frac{1}{t^{3/2}}\right) \exp\left(-\frac{x^2}{2t}\right).\end{aligned}$$

Hence, $\frac{\partial \phi}{\partial t} = \frac{1}{2}\frac{\partial^2 \phi}{\partial x^2}$. ■

We can use Lemma 10.3.2 to give a physical explanation of the connection between Brownian motion and heat diffusion. We define

$$w(t,x) = \mathbb{E}_x[f(B_t)] \tag{10.5}$$

That is, to get $w(t,x)$, we start a particle at location x, let it perform Brownian motion for time t, and then take the expected value of $f(B_t)$.

Lemma 10.3.3 *$w(t,x)$ satisfies the heat equation* (10.3) *and the initial condition* (10.4).

Before we give the proof, let us discuss the physical interpretation of this result. Within our metal rod, the metal atoms have fixed positions. But atoms that are next to each other transfer heat between each other, in random directions. If we could pick on an individual 'piece' of heat and watch it move, it would move like a Brownian motion. Since there are lots of little pieces of heat moving around, and they are *very* small, when we measure temperature we only see the average effect of all the little pieces, corresponding to $\mathbb{E}[\ldots]$.

We should think of the Brownian motion in (10.5) as running in reverse time, so as it tracks (backwards in time) the path through space that a typical piece of heat has followed. Then, after running for time t, it looks at the initial condition to find out how much heat there was initially that its eventual location.

PROOF: We have $B_0 = x$, so

$$w(0,x) = \mathbb{E}_x[f(B_0)] = \mathbb{E}_x[f(x)] = f(x).$$

Hence the initial condition (10.4) is satisfied. We still need to check (10.3). To do so we will allow ourselves to swap $\int$s and partial derivatives[2]. We have

$$\begin{aligned}w(t,x) &= \mathbb{E}_x[f(B_t)] \\ &= \int_{-\infty}^{\infty} f(y)\phi_{t,x}(y)\,dy\end{aligned}$$

so, by Lemma 10.3.2,

$$\begin{aligned}\frac{\partial w}{\partial t} &= \frac{\partial}{\partial t}\int_{-\infty}^{\infty} f(y)\phi_{t,x}(y)\,dy \\ &= \int_{-\infty}^{\infty} f(y)\frac{\partial}{\partial t}\phi_{t,x}(y)\,dy\end{aligned}$$

[2]As far as this course is concerned, there is no need for you to justify interchanging $\int$ and derivatives. In reality, though, it can (occasionally) fail and there are conditions to check. See MAS350/451 for details.

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} f(y) \frac{1}{2} \frac{\partial^2}{\partial x^2} \phi_{t,x}(y)\, dy \\
&= \frac{1}{2} \frac{\partial^2}{\partial x^2} \int_{-\infty}^{\infty} f(y) \phi_{t,x}(y)\, dy \\
&= \frac{1}{2} \frac{\partial^2 w}{\partial x^2}
\end{aligned}
$$

as required. ■

There are many similar ways to connect various stochastic processes to PDEs. This kind of connection can be very useful, because it allows us to transfer knowledge about stochastic process to (and from) knowledge about PDEs. We'll see a more sophisticated example in Section 13.1.

10.4 Properties of Brownian motion

We now examine some of the more detailed properties of Brownian motion. Recall that B_t denotes a standard Brownian motion.

Symmetry

The normal distribution $Z \sim N(0,\sigma^2)$ is symmetric about 0, in the sense that $-Z$ also has the distribution $N(0,\sigma^2)$. This symmetry about 0 is also present in Brownian motion.

Lemma 10.4.1 *The stochastic process $W_t = -B_t$ is a standard Brownian motion.*

PROOF: We must check that $W_t = -B_t$ satisfies the three defining properties of Brownian motion. By the first property, B_t is almost surely continuous, hence $-B_t$ is also almost surely continuous. Also, for $0 \le u \le t$ we have

$$W_t - W_u = -(B_t - B_u).$$

Since, by the second property, $B_t - B_u$ is independent of $\mathcal{F}_u$, so is $W_t - W_u$, and we have

$$\sigma(W_v\,;\,v \le u) = \sigma(-W_v\,;\,v \le u) = \sigma(B_v\,;\,v \le u) = \mathcal{F}_u.$$

Thus $W_t - W_u$ is independent of $\sigma(W_v\,;\,v \le u)$.

Lastly, if $Z \sim N(0,t)$ then, using the symmetry of normal random variables, $-Z \sim N(0,t)$, so we have

$$W_t - W_u = -(B_t - B_u) \sim N(0, t-u)$$

by the third property. Hence, all three properties also hold for (W_t), so W_t is a Brownian motion. Since $W_0 = -B_0 = 0$, we have that (W_t) is a standard Brownian motion. ∎

Lemma 10.4.1 is often referred to as an example of a 'self-symmetry' of Brownian motion, meaning a transformation of a Brownian motion that results in another Brownian motion. It turns out that Brownian motion has many self-symmetries, and they are very important to the theory of Brownian motion.

Non-differentiability

As we've seen in Section 10.1, the paths of Brownian motion look very jagged and erratic. We can express this idea formally: the sample paths of Brownian motion are not differentiable!

Lemma 10.4.2 *Let $t \in [0,\infty)$. Almost surely, the function $t \mapsto B_t$ is not differentiable at t.*

PROOF: Using the second property of Brownian motion, and the scaling properties of normal random variables,

$$\frac{B_{t+h} - B_t}{h} \sim \frac{B_h}{h} \sim \frac{X}{\sqrt{h}}.$$

where $X \sim N(0,1)$. Note that X is positive half the time and negative half the time (and $\mathbb{P}[X = 0] = 0$). Hence, as $h \to 0$, we obtain that

$$\frac{X}{\sqrt{h}} \stackrel{a.s.}{\to} X_\infty = \begin{cases} \infty & \text{with probability } 1/2, \\ -\infty & \text{with probability } 1/2. \end{cases} \tag{10.6}$$

Form Lemma 6.1.2, almost sure convergence implies convergence in distribution, so this limit also holds in distribution. Since $\frac{B_{t+h}-B_t}{h}$ has the same distribution as $\frac{X}{\sqrt{h}}$, we obtain that $\frac{B_{t+h}-B_t}{h}$ converges in distribution to X_∞.

Consider the event $E = \{B_t \text{ is differentiable at } t\}$. When the event E occurs, $\frac{B_{t+h}-B_t}{h}$ converges to a finite quantity as $h \to 0$. However, we saw above that $\frac{B_{t+h}-B_t}{h}$ had the limit X_∞, with $\mathbb{P}[X_\infty \in \{\infty, -\infty\}] = 1$, so the probability that this limit is a finite quantity is zero. Therefore, $\mathbb{P}[E] = 0$. ■

Pure mathematicians discovered functions that were nowhere differentiable at around the start of the 20^{th} century. At first, they were widely thought to be mathematical curiosities, with little or no importance in the 'real' world. A few decades later, the discovery that Brownian motion played a key role in physics, biology and mathematical finance had reversed this viewpoint.

Relationship to martingales

It turns out that there are many martingales associated to Brownian motion. Here's two, with two more to come in exercise **10.6**, and others in later sections of the course.

Lemma 10.4.3 *Brownian motion is a martingale.*

PROOF: It is enough to look at the case (B_t) of standard Brownian motion, since adding and subtracting a deterministic constant does not change if a process is a martingale.

Since $B_t \sim N(0,t)$ we have $\text{var}(B_t) < \infty$, which implies that $B_t \in L^1$. Since the filtration $(\mathcal{F}_t)$ is the generated filtration of B_t, is immediate that B_t is adapted. Lastly, for any $0 \le u \le t$ we have

$$\begin{aligned}
\mathbb{E}[B_t \,|\, \mathcal{F}_u] &= \mathbb{E}[B_t - B_u \,|\, \mathcal{F}_u] + \mathbb{E}[B_u \,|\, \mathcal{F}_u] \\
&= \mathbb{E}[B_t - B_u] + B_u \\
&= B_u.
\end{aligned}$$

Here, we use the properties of Brownian motion: $B_t - B_s$ is independent of $\mathcal{F}_u$ and $\mathbb{E}[B_t] = \mathbb{E}[B_u] = 0$. ■

Lemma 10.4.4 $B_t^2 - t$ *is a martingale*

PROOF: Since $B_t \sim N(0,t)$ we have $\text{var}(B_t) < \infty$, which implies $B_t^2 \in L^1$. Hence also $B_t^2 - t \in L^1$. Since $B_t^2 - t$ is a deterministic function of B_t, we have that $B_t^2 - t$ is adapted to the generated filtration of B_t. Lastly, for $0 \le u \le t$,

$$\begin{aligned}
\mathbb{E}[B_t^2 - t \,|\, \mathcal{F}_u] &= \mathbb{E}[(B_t - B_u)^2 + 2B_tB_u - B_u^2 \,|\, \mathcal{F}_u] - t \\
&= \mathbb{E}[(B_t - B_u)^2 \,|\, \mathcal{F}_u] + 2B_u\mathbb{E}[B_t \,|\, \mathcal{F}_u] - B_u^2 - t \\
&= \mathbb{E}[(B_t - B_u)^2] + 2B_u^2 - B_u^2 - t \\
&= (t-u) + B_u^2 - t \\
&= B_u^2 - u
\end{aligned}$$

as required. Here we use that $B_t - B_u$ is independent of $\mathcal{F}_u$, along with both (10.1) and Lemma 10.4.3. ■

10.5 Exercises on Chapter 10

In all the following questions, B_t denotes Brownian motion.

On Brownian motion

10.1 Consider the process $C_t = \mu t + \sigma B_t$, for $t \geq 0$, where $\mu \in \mathbb{R}$ and $\sigma > 0$ are deterministic constants.

(a) Find the mean and variance of C_t.

(b) Let $0 \leq u \leq t$. What is the distribution of $C_t - C_u$?

(c) Is C_t a random continuous function?

(d) Is C_t a Brownian motion?

10.2 Let $0 \leq u \leq t$. Use the properties of Brownian motion to show that $\mathrm{cov}(B_t, B_u) = u$.

10.3 Let $u \geq 0$ and $t \geq 0$. Show that $\mathbb{E}[B_t \,|\, \mathcal{F}_u] = B_{\min(u,t)}$.

10.4 (a) Show that $\mathbb{E}[B_t^n] = t(n-1)\mathbb{E}[B_t^{n-2}]$ for all $n \geq 2$. *(Hint: Integrate by parts!)*

(b) Deduce that $\mathbb{E}[B_t^2] = t$ and $\mathrm{var}(B_t^2) = 2t^2$.

(c) Write down $\mathbb{E}[B_t^n]$ for any $n \in \mathbb{N}$.

(d) Show that $B_t^n \in L^1$ for all $n \in \mathbb{N}$.

10.5 Let $Z \sim N(\mu, \sigma^2)$. Show that $\mathbb{E}[e^Z] = \exp(\mu + \frac{1}{2}\sigma^2)$. *(Hint: Complete the square!)*

10.6 Show that the following processes are martingales.

(a) $X_t = \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right)$ where $\sigma > 0$ is a deterministic constant.

(b) $Y_t = B_t^3 - 3tB_t$.

10.7 Fix $t > 0$ and for each $n \in \mathbb{N}$ let $(t_k)_{k=0}^n$ be such that $0 = t_0 < t_1 < \ldots < t_n = t$ and $\max_k |t_{k+1} - t_k| \to 0$ as $n \to \infty$. (For example: $t_k = \frac{kt}{n}$).

(a) Show that $\sum\limits_{k=0}^{n-1} t_{k+1} - t_k = t$ and $\sum\limits_{k=0}^{n-1} B_{t_{k+1}} - B_{t_k} = B_t$.

(b) Show that $\sum\limits_{k=0}^{n-1} (t_{k+1} - t_k)^2 \to 0$ as $n \to \infty$.

(c) Set $S_n = \sum\limits_{k=0}^{n-1} (B_{t_{k+1}} - B_{t_k})^2$. Show that $\mathbb{E}[S_n] = t$ and that $\mathrm{var}(S_n) \to 0$ as $n \to \infty$.

Challenge Questions

10.8 (a) Let $y \geq 1$. Show that

$$\mathbb{P}[B_t \geq y] \leq \sqrt{\frac{t}{2\pi}}\, e^{-\frac{y^2}{2t}}.$$

Let $\alpha > \frac{1}{2}$. Deduce that $\mathbb{P}[B_t \geq t^\alpha] \to 0$ as $t \to \infty$. What about $\alpha = \frac{1}{2}$?

(b) Let $y \geq 0$. Show that

$$\mathbb{P}[B_t \geq y] \leq \frac{t}{\sqrt{2\pi}} \frac{1}{y} e^{-\frac{y^2}{2t}}.$$

Deduce that $B_t \to 0$ in probability as $t \searrow 0$.

Chapter 11

Stochastic integration

In this section we introduce stochastic integrals, through the framework of Ito integration. The mathematical framework for stochastic integration was developed in the 1950s, by the Japenese mathematician Kiyoshi Ito (sometimes written Itô). It has grown into becoming one of the most effective modelling tools of the present day.

In Lemma 10.4.2 we showed that Brownian motion was not differentiable. This is awkward, because mathematical modelling often relies on calculus, which (in its classical form) relies heavily on working with derivatives. However, the difficulty can be overcome by forgetting about differentiation and making *integration* the central theme.

11.1 Introduction to Ito calculus

In **classical calculus**, of the sort you are already used to using, we typically deal with objects of the form

$$\int_a^b f(t)\,dt \tag{11.1}$$

where f is a suitably well behaved function. For simplicity, let's take f to be continuous. From an intuitive point of view, we often regard (11.1) as representing the 'area under the curve' f between a and b. This is justified by the fact that we have

$$\int_a^b f(t)\,dt = \lim_{\delta\to 0}\sum_{i=1}^{n} f(t_{i-1})[t_i - t_{i-1}] \tag{11.2}$$

where $(t_i)_{i=0}^n$ is such that $a = t_0 < t_1 < \ldots < t_n = b$ and $\delta = \max_i |t_t - t_{t-1}|$. Note that sending $\delta \to 0$ means that the t_i change position and get closer together, and consequently $n \to \infty$; this is a mild abuse of notation that is commonly used.

Note that if f is a *random* continuous function then (11.1) still makes sense: now, $\int_a^b f(t)\,dt$ is just the area under a random curve; itself a random quantity. This is one way to involve random variables in calculus. There is another:

In **Ito calculus** we are interested in integrals that are written

$$\int_a^b f(t)\,dB_t$$

where B_t is a Brownian motion. Let us begin by discussing what this new type of integral represents; it is *not* the area under a curve.

In (11.2), the dt on the left side corresponds to the $t_i - t_{i-1}$ on the right. By analogy to (11.2), our new dB_t term corresponds to $B_{t_i} - B_{t_{i-1}}$, giving

$$\int_a^b f(t)\,dB_t = \lim_{\delta\to 0}\sum_{i=1}^n f(t_{i-1})[B_{t_i} - B_{t_{i-1}}]. \tag{11.3}$$

Graphically, this means that we measure the widths of the bars using increments of Brownian motion, instead of side-length. For now, let us not worry about which mode of convergence will be used for the limit, or how to choose the t_is.

In order to understand why this is a useful idea, from the point of view of stochastic modelling, we need to think about σ-fields and filtrations. In particular, let us take $f(t)$ to be a stochastic process, and let us assume that $f(t)$ is adapted, with respect to the filtration $\mathcal{F}_t = \sigma(B_s\,;\, s \le t)$. Now, consider the term

$$f(t_{i-1})[B_{t_i} - B_{t_{i-1}}].$$

This formula represents a generic model of taking a decision that then has a random effect. The value of $f(t_{i-1})$ is chosen, based only on information known at time t_{i-1}, then during $t_{i-1} \mapsto t_i$ the world evolves randomly around us, and the effect of our decision combined with this random evolution is represented by $f(t_{i-1})[B_{t_i} - B_{t_{i-1}}]$.

The sum,

$$\sum_{i=1}^n f(t_{i-1})[B_{t_i} - B_{t_{i-1}}] \tag{11.4}$$

corresponds to the cumulative result of multiple decision making steps, at times $t_0 \mapsto t_1 \mapsto t_2 \mapsto \ldots \mapsto t_n$. At each time t_{i-1} a decision is taken for the value of $f(t_i)$, based only on previously available information, then the world changes randomly during $t_{i-1} \mapsto t_i$, and at time t_i we receive and add the random effect of our decision: $f(t_{i-1})[B_{t_i} - B_{t_{i-1}}]$.

Remark 11.1.1 We've seen this idea before, in Section 7.2 when we modelled roulette using the martingale transform. If we set $t_i = i$, $C_n = f(t_n)$ and $M_n = B_{t_n} = B_n$, then M_n is a discrete time martingale (take $\mathcal{F}_n = \sigma(B_i\,;\, i \le n)$), and (11.4) is precisely the martingale transform $(C \circ M)_n = \sum_{i=1}^n C_{i-1}(M_i - M_{i-1})$.

The final stage of this intuition is to understand the limit in (11.3). Now we take decisions at times

$$a = t_0 \mapsto t_1 \mapsto t_2 \ldots \mapsto t_n = b$$

as $\delta = \max_i |t_i - t_{i-1}| \to 0$. The corresponds to taking a *continuous* stream of decisions during the time interval $[a, b]$, each based on previously available information, each of which has an (infinitesimally) small effect. The stochastic integral,

$$\int_a^b f(t)\,dB_t$$

corresponds to the cumulative effect of all these decisions.

Of course, the situation we are most interested in, within this course, is that of managing a portfolio. In continuous time we can *continually* take decisions to buy and sell based on the information that is currently available to us. Our $f(t_i)$ will be a process relating to the stocks that we hold, and the Brownian motion B_t will provide the randomness that moves stock prices up and down. Developing the details of this modelling effort, and the pricing results that come out of it, will take up *the rest of the course.*

Remark 11.1.2 In this section it was helpful to write the integrand as $f(t)$. Since the integrand is a a stochastic process, we will often (but by no means always) stick to our convention of denoting stochastic processes with capital letters, such as F_t, giving $\int_a^b F_t\,dB_t$.

Remark 11.1.3 As an alternative approach to defining the meaning of Ito integrals it might be tempting to try and write

$$\int_a^b F_t\,dB_t = \int_a^b F_t\,\frac{dB_t}{dt}\,dt$$

and use this idea to relate stochastic integrals to classical integrals. Unfortunately, the right hand side of the above expression does not make sense - we have shown in Lemma 10.4.2 that B_t is not differentiable, so $\frac{dB_t}{dt}$ does not exist.

11.2 Ito integrals

In Section 11.1 we discussed the ideas behind Ito integrals. We did not discuss one key (theoretical) question: if and when the limit in (11.3) actually exists?

Let us recall our usual notation. We work over a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, where the filtration $\mathcal{F}_t$ is the generated filtration of a Brownian motion B_t. We use the letters t, u and sometimes also v, as our time variables.

We say that a stochastic process F_t is *locally square integrable* if

$$\int_0^t \mathbb{E}\left[F_u^2\right] du < \infty. \tag{11.5}$$

for all $t \in [0, \infty)$. We define $\mathcal{H}^2$ to be the set of locally square integrable continuous stochastic processes $F = (F_t)_{t=0}^{\infty}$ that are adapted to $(\mathcal{F}_t)$.

It turns out that the condition $F \in \mathcal{H}^2$ is the correct condition under which to take the limits discussed in Section 11.1. The following theorem formally states that Ito integrals exist, and gives some of their first properties.

Theorem 11.2.1 *For any $F \in \mathcal{H}^2$, and any $t \in [0, \infty)$ the Ito integral*

$$\int_0^t F_t \, dB_t$$

exists, and is a continuous martingale with mean and variance given by

$$\mathbb{E}\left[\int_0^t F_t \, dB_t\right] = 0,$$
$$\mathbb{E}\left[\left(\int_0^t F_t \, dB_t\right)^2\right] = \int_0^t \mathbb{E}[F_t^2] \, dt.$$

So far we have only looked at integrals over $[0, t]$. We can extend the definition to $\int_a^b$, simply by repeating the whole procedure above with limits $[a, b]$ instead of $[0, t]$. It is easily seen that this gives the usual consistency property

$$\int_a^c F_t \, dB_t = \int_a^b F_t \, dB_t + \int_b^c F_t \, dB_t \tag{11.6}$$

for $a \le b \le c$. We won't include a proof of this in our course.

Like classical integrals, Ito integrals are linear. For $\alpha, \beta \in \mathbb{R}$ we have

$$\int_a^b \alpha F_t + \beta G_t \, dB_t = \alpha \int_a^b F_t \, dB_t + \beta \int_a^b G_t \, dB_t. \tag{11.7}$$

Again, we won't include a proof of this formula in our course.

In future, we'll use the linearity and consistency properties without comment. However, as we'll explore in the next two sections, there are many ways in which the Ito integral does *not* behave like the classical integral.

Comparing Ito integration to classical integration

Let us first note one similarity. It is true that

$$\int_0^t 0 \, dB_u = 0.$$

This matches classical integrals, where we have $\int_0^t 0\,du = 0$. We can see this from 11.3, by setting $f \equiv 0$, and noting that the limit of 0 is 0.

Here's a first difference: fix some $t > 0$ and let us look at $\int_0^t 1\,dB_u$. If we set $f \equiv 1$ in (11.3), then we obtain $\sum_{i=1}^n f(t_{i-1})[B_{t_i} - B_{t_{i-1}}] = B_t - B_0 = B_t$, and hence

$$\int_0^t 1\,dB_u = B_t. \tag{11.8}$$

Of course, in classical calculus we have $\int_0^t 1\,du = t$. This is our first example of an important principle: Ito integration behaves differently to classical integration. To illustrate further, in Section 12.1 we will put together a set of tools for calculating Ito integrals, and in Example 12.1.2 we will see that

$$\int_0^t B_u\,dB_u = \frac{B_t^2}{2} - \frac{t}{2}.$$

This corresponds to taking f to be the identity function in (11.3), $f(x) = x$. Of course, in classical calculus we have $\int_0^t u\,du = \frac{u^2}{2}$, which is very different.

11.3 Existence of Ito integrals ($\star$)

This section is off-syllabus, and as such is marked with ($\star$). It will not be covered in lectures.

The argument that proves Theorem 11.2.1, through justifying the limit taken in (11.3), is based heavily on martingales, metric spaces and Hilbert spaces. It comes in two steps, the first of which involves a class of stochastic processes F known as simple processes – see Definition 11.3.1 below. The second step uses limits extends the definition for simple processes onto a much larger class. We'll look at these two steps in turn.

We'll use the notation $\wedge$ and $\vee$ from Chapter 8. That is, we write $\min(s,t) = s \wedge t$ and $\max(s,t) = s \vee t$.

Definition 11.3.1 We say that a stochastic process F_u is a simple process if there exists deterministic points in time $0 = t_0 < t_1 < \ldots < t_m$ such that:

1. F_u remains constant during each interval $u \in [t_{t-1}, t_i)$, and $F_u = 0$ for $u \geq t_m$.
2. For each i, F_{t_i} is bounded and $F_{t_i} \in \mathcal{F}_{t_i}$.

For a simple process F, with (t_i) as in Definition 11.3.1 we define

$$I_F(t) = \sum_{i=1}^{n} F_{t_{i-1}}[B_{t_i \wedge t} - B_{t_{i-1} \wedge t}]. \tag{11.9}$$

Note that this is essentially the right hand side of (11.3) but without the limit. The point of the $\wedge t$ is that we are aiming to define an integral over $[0,t]$; the $\wedge t$ makes sure that $I_F(t)$ only picks up increments from the Brownian motion during $[0,t]$.

We can already see the connection to martingales (which builds on Remark 11.1.1):

Lemma 11.3.2 *Suppose that F_t is a simple process. Then $I_F(t)$ is an $\mathcal{F}_t$ martingale.*

PROOF: Since a (finite) sum of martingales is a also martingale, it is enough to fix i and show that $M_t = F_{t_{i-1}}[B_{t_i \wedge t} - B_{t_{i-1} \wedge t}]$ is a martingale. The argument is rather messy, because we have to handle the $\wedge t$ everywhere.

Let us look first at L^1. F_{t_i} is bounded we have some deterministic $A \in \mathbb{R}$ such that $|F_{t_i}| \leq A$ (almost surely). Hence, $\mathbb{E}\left[|F_{t_{i-1}}[B_{t_i \wedge t} - B_{t_{i-1} \wedge t}]|\right] \leq A\,\mathbb{E}[|B_{t_i \wedge t} - B_{t_{i-1} \wedge t}|] < \infty$. Here, we use that $B_{t_i \wedge t} - B_{t_{i-1} \wedge t} \sim N(0,\, t_i \wedge t - t_{i-1} \wedge t)$, which is in L^1. Hence, $M_t \in L^1$.

Next, adaptedness, for which we consider two cases.

- If $t \geq t_{i-1}$ then $F_{t_{i-1}} \in \mathcal{F}_t$. Since $t_i \wedge t \leq t$, we have $B_{t_i \wedge t} \in m\mathcal{F}_t$ and, similarly, $B_{t_{i-1} \wedge t} \in m\mathcal{F}_t$, hence also $M_t \in m\mathcal{F}_t$.
- If $t < t_{i-1}$ then $t_i \wedge t = t_{i-1} \wedge t = t$, meaning that $B_{t_i \wedge t} - B_{t_{i-1} \wedge t} = 0$. So $M^{(i)}_{t_i} = 0$, which is deterministic and therefore also in $m\mathcal{F}_t$.

Therefore, (M_t) is adapted to $(\mathcal{F}_t)$.

Lastly, let $0 \leq u \leq t$. Again, we consider two cases.

- If $u \geq t_{i-1}$ then $F_{t_{i-1}} \in \mathcal{F}_u$ and we have

$$\mathbb{E}\left[F_{t_{i-1}}[B_{t_i \wedge t} - B_{t_{i-1} \wedge t}] \,|\, \mathcal{F}_u\right] = F_{t_{i-1}}\left(\mathbb{E}\left[B_{t_i \wedge t} \,|\, \mathcal{F}_u\right] - \mathbb{E}\left[B_{t_{i-1} \wedge t} \,|\, \mathcal{F}_u\right]\right)$$

$$
\begin{aligned}
&= F_{t_{i-1}}[B_{t_i\wedge t\wedge u} - B_{t_{i-1}\wedge t\wedge u}] \\
&= F_{t_{i-1}}[B_{t_i\wedge u} - B_{t_{i-1}\wedge u}].
\end{aligned}
$$

Here, in the first line we take out what is known, and we use the martingale property of Brownian motion to deduce the second line. The third line then follows because $u \le t$.

- If $u < t_{i-1}$ then $B_{t_i\wedge u} - B_{t_{i-1}\wedge u} = 0$. Also, by the tower rule

$$
\begin{aligned}
\mathbb{E}\left[F_{t_{i-1}}[B_{t_i\wedge t} - B_{t_{i-1}\wedge t}]\,|\,\mathcal{F}_u\right] &= \mathbb{E}\left[\mathbb{E}\left[F_{t_{i-1}}[B_{t_i\wedge t} - B_{t_{i-1}\wedge t}]\,|\,\mathcal{F}_{t_{i-1}}\right]\,|\,\mathcal{F}_u\right] \\
&= \mathbb{E}\left[F_{t_{i-1}}\left(\mathbb{E}\left[B_{t_i\wedge t}\,|\,\mathcal{F}_{t_{i-1}}\right] - \mathbb{E}\left[B_{t_{i-1}\wedge t}\,|\,\mathcal{F}_{t_{i-1}}\right]\right)\,|\,\mathcal{F}_u\right] \\
&= \mathbb{E}\left[F_{t_{i-1}}\left(B_{t_i\wedge t\wedge t_{i-1}} - B_{t_{i-1}\wedge t\wedge t_{i-1}}\right)\,|\,\mathcal{F}_u\right] \\
&= \mathbb{E}\left[F_{t_{i-1}}\left(B_{t_{i-1}} - B_{t_{i-1}}\right)\,|\,\mathcal{F}_u\right] \\
&= 0.
\end{aligned}
$$

In both cases, we have shown that $\mathbb{E}[M_t\,|\,\mathcal{F}_u] = M_u$. ■

Lemma 11.3.3 *Suppose that F_t is a simple process. Then, for any $0 \le t \le \infty$,*

$$\mathbb{E}\left[I_F(t)^2\right] = \int_0^t \mathbb{E}\left[F_u^2\right]\,du. \tag{11.10}$$

PROOF: See exercise 11.10. The proof similar in style to that of Lemma 11.3.2. ■

Essentially, Theorem 11.2.1 says that Ito integrals exist for $F \in \mathcal{L}^2$ and that Lemmas 11.3.2 and 11.3.3 are true, not just for simple processes, but for Ito integrals in general. This observation brings us to second step of the construction of Ito integrals, although we won't be able to cover all of the details here. It comes in two sub-steps:

1. Fix $t < \infty$ and begin with a process $F \in \mathcal{H}^2$. Approximate F by a sequence of simple processes $F^{(k)}$ such that
$$\int_0^t \mathbb{E}\left[\left(F_u - F_u^{(k)}\right)^2\right]\,du \to 0 \tag{11.11}$$
as $k \to \infty$. It can be proved that this is always possible.

2. For each k, $I_{F_m}(t)$ is defined by (11.9). We define
$$\int_0^t F_u\,dB_u = \lim_{k\to\infty} I_{F^{(k)}}(t). \tag{11.12}$$
Using (11.11), it can be shown that this limit exists, with convergence in L^2, and moreover its value (on the left hand side) is independent of the choice of approximating sequence $F^{(k)}$ (on the right hand side).

We end with a brief summary of the mathematics that lies behind (11.11) and (11.12). We have shown that the map $F \mapsto I_F$ takes a sample process, which is an example of a locally square integrable adapted stochastic process, and gives back a martingale that is in $\mathcal{L}^2$. If we add appropriate restrictions on the left and right continuity of F, it can be shown that the map $F \mapsto I_F$ becomes a linear operator between two Hilbert spaces. Further, (11.10) turns out to be precisely the statement that $F \mapsto I_F$ is an isometry (usually referred to as the *Ito isometry*). The set of simple stochastic processes is a dense subset of the space of square integrable adapted stochastic processes, which allows us to use a powerful theorem about isometries between Hilbert spaces (known as the completion theorem) to take the limit in (11.12).

11.4 Ito processes

We are now ready to define precisely the types of stochastic process that we will be interested in for most of the remainder of this course.

Definition 11.4.1 A stochastic process X is known as an Ito process if X_0 is $\mathcal{F}_0$ measurable and X can be written in the form

$$X_t = X_0 + \int_0^t F_u\,du + \int_0^t G_u\,dB_u \tag{11.13}$$

Here, $G \in \mathcal{H}^2$ and F is a continuous adapted process.

The first integral is a classical integral: the area under the random curve F_t. The second integral is an Ito integral.

Note that, to fully specify X_t, we also need to know both F_u, G_u, and the initial value X_0. Since this means we'll be dealing with integrals of the form $\int_0^t F_u\,du$, it is helpful for us to know a fact from integration theory:

Lemma 11.4.2 *For a continuous stochastic process F, if one (which $\Rightarrow$ both) of the two sides is finite, then we have $\mathbb{E}\left[\int_0^t F_u\,du\right] = \int_0^t \mathbb{E}[F_u]\,du$.*

In words, we can swap $\int\,du$ and $\mathbb{E}$s as long as we aren't dealing with ∞ (warning: it doesn't work for Ito integrals, see exercise **11.8**!). We won't include a proof in this course, but you can find one which works for both $\int$s and $\sum$s in MAS451/6352.

We can calculate the expectation of X_t using Lemma 11.4.2.

$$\begin{aligned}\mathbb{E}[X_t] &= \mathbb{E}[X_0] + \mathbb{E}\left[\int_0^t F_u\,du\right] + \mathbb{E}\left[\int_0^t G_u\,dB_u\right].\\ &= \mathbb{E}[X_0] + \int_0^t \mathbb{E}\left[F_u\right]\,du.\end{aligned}$$

Here, Lemma 11.4.2 allows us to swap $\int$ and $\mathbb{E}$ for the du integral. The term $\mathbb{E}[\int_0^t G_u\,dB_u]$ is zero because Theorem 11.2.1 told us that $\int_0^t G_u\,dB_u$ was martingale with zero mean.

Example 11.4.3 Let X_t be the Ito process satisfying

$$X_t = 1 + \int_0^t 2B_u^2\,du + \int_0^t 3u\,dB_u.$$

Then

$$\begin{aligned}\mathbb{E}[X_t] &= \mathbb{E}[1] + \mathbb{E}\left[\int_0^t 2B_u^2\,du\right] + \mathbb{E}\left[\int_0^t 3u\,dB_u\right]\\ &= 1 + 2\int_0^t \mathbb{E}[B_u^2]\,du\\ &= 1 + 2\int_0^t u\,du\\ &= 1 + t^2.\end{aligned}$$

Example 11.4.4 Brownian motion is an Ito process. It satisfies $B_t = B_0 + \int_0^t 0\,du + \int_0^t 1\,dB_u$.

Equating coefficients

A useful fact about Ito processes is that we can 'equate coefficients' in much the same way as we equate the coefficients of terms in polynomials. To be precise, we have

Lemma 11.4.5 *Suppose that*

$$X_t = X_0 + \int_0^t F_u^X \, du + \int_0^t G_u^X \, dB_u$$
$$Y_t = Y_0 + \int_0^t F_u^Y \, du + \int_0^t G_u^Y \, dB_u$$

are Ito processes and that $\mathbb{P}[\text{for all } t, X_t = Y_t] = 1$. *Then,*

$$\mathbb{P}[\text{for all } t, F_t^X = F_t^Y \text{ and } G^X = G_t^Y] = 1.$$

Proof of this lemma is outside of the scope of our course. However, the result will be very important to us, since it is key to the argument that will allow to hedge financial derivatives in continuous time.

11.5 Exercises on Chapter 11

In all the following questions, B_t denotes a Brownian motion and $\mathcal{F}_t$ denotes its generated filtration.

On Ito integration

11.1 Using (11.8), find $\int_v^t 1\,dB_u$, where $0 \leq v \leq t$.

11.2 Show that the process e^{B_t} is in $\mathcal{H}^2$. (Hint: Use (10.2).)

11.3 (a) Let $Z \sim N(0,1)$. Show that the expectation of $e^{\frac{Z^2}{2}}$ is infinite.

(b) Give an example of a continuous, adapted, stochastic process that is not in $\mathcal{H}^2$.

11.4 Let X_t be an Ito process satisfying

$$X_t = 2 + \int_0^t t + B_u^2\,du + \int_0^t B_u^2\,dB_u.$$

Find $\mathbb{E}[X_t]$.

11.5 Which of the following stochastic processes are Ito processes?

(a) $X_t = 0$,

(b) $Y_t = t^2 + B_t$,

(c) The symmetric random walk from Section 4.1.

11.6 Let V_t be the stochastic process given by

$$V_t = e^{-kt}v + \sigma e^{-kt}\int_0^t e^{ku}\,dB_u$$

where $k, \sigma, v > 0$ are deterministic constants. Find the mean and variance of V_t.

11.7 Suppose that $\mu > 0$ is a deterministic constant and that $\sigma_t \in \mathcal{H}^2$. Let X_t be given by

$$X_t = \int_0^t \mu\,du + \int_0^t \sigma\,dB_u.$$

Show that X_t is a submartingale.

11.8 (a) Give an example of a stochastic process F_t such that $\int_0^t \mathbb{E}[F_s]\,ds = \mathbb{E}[\int_0^t F_s\,dB_s]$.

(b) Give an example of a stochastic process F_t such that $\int_0^t \mathbb{E}[F_s]\,ds \neq \mathbb{E}[\int_0^t F_s\,dB_s]$.

Challenge Questions

11.9 (a) Let X and Y be random variables in L^2. Show that

$$2|\mathbb{E}[XY]| \leq \mathbb{E}[X^2] + \mathbb{E}[Y^2]$$

(b) Show that $\mathcal{H}^2$ is a real vector space.

11.10 ($\star$) Prove Lemma 11.3.3.

Chapter 12

Stochastic differential equations

The situation we have arrived at is that we know Ito integrals exist but, as yet, we are unable to calculate them or do much calculation with them. We will address this issue in Section 12.1 but first, in order to make our calculations run smoothly, we need to introduce some new notation.

We now understand how to make sense of equations of the form

$$X_t = X_0 + \int_0^t F_u \, du + \int_0^t G_u \, dB_t \tag{12.1}$$

where B_t is Brownian motion. Note that we allow cases in which the stochastic processes F_u, G_u depend on X_u (for example, we could have $F_u = X_t^2$), but that X_t is an unknown stochastic process. Equations of this type are known as *stochastic differential equations*, or SDEs for short – an unfortunate name because they have no differentiation involved! They are also sometimes known as stochastic integral equations, but for historical reasons the term SDE has become the most commonly used.

'Solving' the equation (12.1) essentially means finding X_t in terms of B_t. If F_t and G_t depend only on t and B_t, then (12.1) is just an explicit formula, which automatically that tells us that there is a solution to (11.13). However, if F_t and/or G_t depend on X_t (e.g. $F_t = 2X_t$) then (12.1) is not an explicit formula and there is no guarantee that a solution for X_t exists.

Remark 12.0.1 ($\star$) The theory of existence and uniqueness of solutions to SDEs relies on analysis in more delicate ways than we have time to discuss in this course. We use the term 'solution' for what is usually referred to in the theory of SDEs as a 'strong solution'.

In general SDEs, like their classical counterparts ODEs, often do not have explicit solutions, and frequently have no solutions. Happily, though, in all the cases we need to consider, we will be able to write down explicit solutions.

Writing $\int$s everywhere is cumbersome, so it is common to 'drop the $\int$s' and write (12.1) as

$$dX_t = F_t \, dt + G_t \, dB_t \tag{12.2}$$

This equation has exactly the same meaning as (12.1), it is just written in different notation (to be clear: we are not differentiating anything). The notation dX_t, dB_t used in (12.2) is known as the notation of *stochastic differentials*, and we'll use it from now on.

When we convert from stochastic differential form (12.2) to integral form (12.1) we can choose which limits to put onto the integrals. In (12.1) we choose $[0, t]$, but if $v \le t$ then we can also

choose $[v,t]$, giving
$$X_t = X_v + \int_v^t F_u\,du + \int_v^t G_u\,dB_u.$$
(Rigorously, we can do this because (12.1) implies (12.2) with v in place of t, which we can then subtract from (12.2) to obtain limits $[v,t]$.)

We can rewrite our definition of an Ito process in our new notation.

Definition 12.0.2 A stochastic process X_t is an Ito process if it satisfies
$$dX_t = F_t\,dt + G_t\,dB_t$$
for some $G \in \mathcal{H}^2$ and a continuous adapted stochastic process F.

We need one more piece of notation. Given an Ito process X_t, as in Definition 12.0.2, and a stochastic process H_t, we will often write
$$dZ_t = H_t\,dX_t \tag{12.3}$$
which (as a definition) we interpret to mean that
$$\begin{aligned} dZ_t &= H_t(F_t\,dt + G_t\,dB_t) \\ &= H_tF_t\,dt + H_tG_t\,dB_t. \end{aligned}$$
In integral form this represents
$$Z_t = Z_0 + \int_0^t H_uF_u\,du + \int_0^t H_uG_u\,dB_u.$$
Of course, it is much neater to write (12.3).

Remark 12.0.3 ($\star$) There is a limiting procedure that can extend the Ito integral, for a suitable class of stochastic processes Z, to define $\int_0^t H_u\,dZ_u$ directly: in similar style to (11.12) but with increments of Z_t in place of increments of B_t. This approach relies on some difficult analysis, and we won't discuss it in this course.

12.1 Ito's formula

We have commented that, whilst we do know that Ito integrals exist, we are not yet able to do any serious calculations with them. In fact, the situation is similar to that of conditional expectation: direct calculation is usually difficult and, instead, we prefer to work with Ito integrals via a set of useful properties. If you think about it, this is also the situation in classical calculus – you rely on the chain and product rules, integration by parts, etc.

From 11.8, we already know that there are differences between Ito calculus and classical calculus. In fact, there is bad news: *none* of the usual rules[1] used of classical calculus hold in Ito calculus.

There is also good news: Ito calculus does have its own version of the chain rule, which is known as *Ito's formula*. Perhaps surprisingly, this alone turns out to be enough for most purposes[2].

As in Definition 11.4.1, let X be an Ito process satisfying

$$dX_t = F_t\,dt + G_t\,dB_t \tag{12.4}$$

where $G \in \mathcal{H}^2$ and F is a continuous adapted process.

Lemma 12.1.1 (Ito's formula) *Suppose that, for $t \in \mathbb{R}$ and $x \in \mathbb{R}$, $f(t,x)$ is a deterministic function that is differentiable in t and twice differentiable in x. Then $Z_t = f(t, X_t)$ is an Ito process and*

$$dZ_t = \left\{\frac{\partial f}{\partial t}(t,X_t) + F_t\frac{\partial f}{\partial x}(t,X_t) + \frac{1}{2}G_t^2\frac{\partial^2 f}{\partial x^2}(t,X_t)\right\}dt + G_t\frac{\partial f}{\partial x}(t,X_t)\,dB_t.$$

As in classical calculus, it is common to suppress the arguments (t, X_t) of f and its derivatives. This results in simply

$$dZ_t = \left\{\frac{\partial f}{\partial t} + F_t\frac{\partial f}{\partial x} + \frac{1}{2}G_t^2\frac{\partial^2 f}{\partial x^2}\right\}dt + G_t\frac{\partial f}{\partial x}\,dB_t. \tag{12.5}$$

which is the notation we'll usually use. It is sometimes helpful to simplify the expression even further, using (12.3) and (12.4), to $dZ_t = \{\frac{\partial f}{\partial t} + \frac{1}{2}G_t^2\frac{\partial^2 f}{\partial x^2}\}\,dt + \frac{\partial f}{\partial x}\,dX_t$.

Before we say a few words about the proof, let us practice using Ito's formula. We will need it repeatedly throughout the whole of remainder of the course.

Example 12.1.2 Let us apply Ito's formula to calculate dZ_t where $Z_t = B_t^2$. We have $Z_t = f(t, B_t)$ where $f(t,x) = x^2$, which gives $\frac{\partial f}{\partial t} = 0$, $\frac{\partial f}{\partial x} = 2x$ and $\frac{\partial^2 f}{\partial x^2} = 2$. We'll also use that B_t is an Ito process satisfying $dB_t = 0\,dt + 1\,dB_t$. From Ito's formula,

$$\begin{aligned} dZ_t &= \left(0 + (0)(2B_t) + \frac{1}{2}(1^2)(2)\right)dt + (1)(2B_t)\,dB_t \\ &= 1\,dt + 2B_t\,dB_t. \end{aligned}$$

In integral form this reads

$$Z_t = Z_0 + \int_0^t 1\,du + \int_0^t 2B_u\,dB_u.$$

[1] Meaning: the chain rule, product rule, quotient rule, integration by parts, inverse function rule, substitution rule, implicit differentiation rule, ...

[2] Actually, Ito calculus has a product rule too, but we won't need it.

Rearranging, we obtain that

$$\int_0^t B_u\,dB_u = \frac{B_t^2}{2} - \frac{t}{2}. \tag{12.6}$$

This shows that Ito calculus behaves very differently to classical calculus (of course, $\int_0^t u\,du = \frac{u^2}{2}$).

Example 12.1.3 Suppose that X satisfies $dX_t = \mu\,dt + \sigma\,dB_t$, where $\mu \in \mathbb{R}$ and $\sigma > 0$ are deterministic constants. Let $Z_t = X_t e^t$. We want to find dZ_t.

We have $Z_t = f(t, X_t)$ where $f(t,x) = xe^t$, and $F_t = \mu$, $G_t = \sigma$ so by Ito's formula,

$$\begin{aligned} dZ_t &= \left(X_t e^t + (\mu)(e^t) + \frac{1}{2}(\sigma^2)(0)\right)dt + (\sigma)(e^t)\,dB_t \\ &= (X_t + \mu)\,e^t\,dt + \sigma e^t\,dB_t. \end{aligned}$$

Example 12.1.4 Suppose that we want to calculate $\mathbb{E}[B_t^4]$.

We define $Z_t = B_t^4$ and use Ito's formula to find dZ_t. We have $Z_t = f(t, B_t)$ where $f(t,x) = x^4$. Note that Brownian motion B_t is an Ito process, with $dB_t = 0\,dt + 1\,dB_t$. So,

$$\begin{aligned} dZ_t &= \left(0 + (0)(4B_t^3) + \frac{1}{2}(1^2)(12B_t^2)\right)dt + (1)(4B_t^3)\,dB_t \\ &= 6B_t^2\,dt + 4B_t^3\,dB_t. \end{aligned}$$

Hence, in integral form,

$$Z_t = Z_0 + \int_0^t 6B_u^2\,du + \int_0^t 4B_u^3\,dB_u.$$

Taking expectations, and noting that $Z_0 = B_0^4 = 0$,

$$\begin{aligned} \mathbb{E}[Z_t] &= \mathbb{E}\left[\int_0^t 6B_u^2\,du\right] + \mathbb{E}\left[\int_0^t 4B_u^3\,dB_u\right] \\ &= \int_0^t \mathbb{E}\left[6B_u^2\right]\,du + 0 \\ &= \int_0^t 6u\,du \\ &= \frac{6t^2}{2} \\ &= 3t^2. \end{aligned}$$

Here, we used Lemma 11.4.2 to swap $\int$ and $\mathbb{E}$ for the du integral, and to deduce the second line we recall from Theorem 11.2.1 that Ito integrals $\int_0^t \ldots\,dB_t$ are martingales with mean zero.

The result we have obtained matches that from exercise 10.4, but with much less work!

Sketch proof of Ito's formula ($\star$)

The proof of Ito's formula is very technical, and even some advanced textbooks on stochastic calculus omit a full proof. We are now getting used to the principle that (in continuous time) proofs of most important results about stochastic processes make heavy use of analysis; in this case, Taylor's theorem. We'll attempt to give just an indication of where (12.5) comes from.

Fix an interval $[0, t]$ and take t_k such that

$$0 = t_0 < t_1 < t_2 < \ldots < t_n = t.$$

We plan eventually to take a limit as $n \to \infty$, where the minimal distance between two neighbouring t_k goes to zero. Note that this in similar style to the limit used in the construction of Ito integrals. We'll use the notation (just in this section)

$$\begin{aligned}
\Delta t &= t_{k+1} - t_k \\
\Delta B &= B_{t_{k+1}} - B_{t_k} \\
\Delta X &= X_{t_{k+1}} - X_{t_k}.
\end{aligned}$$

We begin by writing

$$f(t, X_t) - f(0, X_0) = \sum_{k=0}^{n-1} f(t_{k+1}, X_{t_{k+1}}) - f(t_k, X_{t_k}).$$

Then, we apply the two dimensional version of Taylor's Theorem to f on the time interval $[t_k, t_{k+1}]$ to give us

$$\begin{aligned}
f(t_{k+1}, X_{t_{k+1}}) - f(t_k, X_{t_k}) = \frac{\partial f}{\partial t}\Delta t + \frac{\partial f}{\partial x}\Delta X + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\Delta X)^2 + \frac{\partial^2 f}{\partial x^2}\Delta X \Delta t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\Delta t)^2 \\
+ [\text{higher order terms}] \qquad (12.7)
\end{aligned}$$

We suppress the argument (t_k, X_{t_k}) of all partial derivatives of f. In the 'higher order terms' we have terms containing $(\Delta X)^3, \Delta t(\Delta X)^2$ and so on. Using the SDE (12.4) we have

$$\begin{aligned}
\Delta X = X_{t_{k+1}} - X_{t_k} &= \int_{t_k}^{t_{k+1}} F_u \, du + \int_{t_k}^{t_{k+1}} G_u \, dB_u \\
&\approx F_{t_k}\Delta t + G_{t_k}\Delta B.
\end{aligned}$$

Summing (12.7) over $\sum_k := \sum_{k=0}^{n-1}$ and using this approximation, we have

$$f(t, X_t) - f(0, X_0) = I_1 + I_2 + I_3 + J_1 + J_2 + J_3 + [\text{higher order terms}]$$

where

$$
\begin{aligned}
I_1 &= \sum_k \frac{\partial f}{\partial t}\Delta t && \to \int_0^t \frac{\partial f}{\partial x}\,du \\
I_2 &= \sum_k \frac{\partial f}{\partial x}F_{t_k}\Delta t && \to \int_0^t \frac{\partial f}{\partial x}F_u\,du \\
I_3 &= \sum_k \frac{\partial f}{\partial x}G_{t_k}\Delta B && \to \int_0^t \frac{\partial f}{\partial x}G_u\,dB_u \\
J_1 &= \frac{1}{2}\sum_k \frac{\partial^2 f}{\partial x^2}G_{t_k}^2(\Delta B)^2 && \to \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}G_u^2\,du \\
J_2 &= \sum_k \frac{\partial^2 f}{\partial x^2}F_{t_k}G_{t_k}(\Delta B)(\Delta t) && \to 0 \\
J_3 &= \sum_k \frac{\partial^2 f}{\partial x^2}F_{t_k}^2(\Delta t)^2 && \to 0
\end{aligned}
$$

As we let $n \to \infty$, and the t_k become closer together, $\Delta t \to 0$ and the convergence shown takes place. In the case of I_1 and I_2 this is essentially by definition of the (classical) integral. For I_3, it is by the definition of the Ito integral, as in (11.3). For J_1, J_2 and J_3 the picture is more complicated; convergence in this case follows by an extension of exercise **10.7**. Essentially, exercise **10.7** tells us that terms of order Δt matter and that $(\Delta B)^2 \approx \Delta t$, resulting in

$$
J_1 \approx \frac{1}{2}\sum_k \frac{\partial^2 f}{\partial x^2}G_{t_k}^2\Delta t \qquad \to \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}G_u^2\,du.
$$

However, $(\Delta t)^2$ and $(\Delta t)(\Delta B) \approx (\Delta t)^{3/2}$ are both much smaller than Δt, with the result that the terms J_2 and J_3 vanish as $\Delta t \to 0$. The higher order terms in (12.7) also vanish. Providing rigorous arguments to take all these limits is the bulk of the work involved in a full proof of Ito's formula.

After the limit has been taken, Ito's formula is obtained by collecting the non-zero terms I_1, I_2, I_3, J_1 together and writing the result in the notation of stochastic differentials.

12.2 Geometric Brownian motion

In this section we focus on a particular SDE, namely

$$dX_t = \alpha X_t \, dt + \sigma X_t \, dB_t \tag{12.8}$$

where $\alpha \in \mathbb{R}$ and $\sigma \geq 0$ are deterministic constants. The parameter α is known as the *drift*, and σ is known as the *volatility*. Equation (12.8) will be important to us because it will be our next step in establishing better models for stock prices (to be continued, in Section 14.1).

The solution to equation (12.8), which we will shortly show exists, is known as *geometric Brownian motion*.

The key step to solving (12.8) is to work with the logarithm of X. With this aim in mind, we *assume* (for now) that there is a solution X that is strictly positive, and consider the process

$$Z_t = \log X_t.$$

Of course our assumption may not be true - but if such a solution does exist we hope to (do some calculations and with Z and) find an explicit formula for it, at which point we can go back and check we really do have a solution.

Remark 12.2.1 Taking logarithms is a natural idea to try. To see this, consider the special case $\sigma = 0$, where we find ourselves back in the world of differential equations: $x(t) = \int_0^t \alpha x(u) \, du$. The fundamental theorem of calculus gives $\frac{dx}{dt} = \alpha x$. We can solve this equation by considering $z = \log x$, obtaining

$$\frac{dz}{dt} = \frac{dz}{dx}\frac{dx}{dt} = \frac{1}{x}\alpha x = \alpha$$

Thus $z(t) = \alpha t + C$ and $x(t) = e^{z(t)} = C' e^{\alpha t}$.

Using Ito's formula, with $Z_t = \log X_t$ (i.e. $f(t,x) = \log x$) we obtain

$$\begin{aligned} dZ_t &= \left(0 + \alpha X_t \frac{1}{X_t} + \frac{1}{2}(\sigma X_t)^2 \frac{-1}{X_t^2}\right) dt + \sigma X_t \frac{1}{X_t} \, dB_t \\ &= \left(\alpha - \frac{1}{2}\sigma^2\right) dt + \sigma \, dB_t. \end{aligned}$$

In integral form this gives us

$$\begin{aligned} Z_t &= Z_0 + \int_0^t \alpha - \frac{1}{2}\sigma^2 \, du + \int_0^t \sigma \, dB_t \\ &= Z_0 + \left(\alpha - \frac{1}{2}\sigma^2\right) t + \sigma B_t. \end{aligned}$$

Since $Z_t = \log X_t$, raising both sides to the power e gives us

$$X_t = X_0 \exp\left((\alpha - \tfrac{1}{2}\sigma^2)t + \sigma B_t\right). \tag{12.9}$$

As we said, this was all based on the assumption that a (strictly positive) solution exists. But, now we have found the formula (12.9), we can go back and check that it does really give us a solution. This part is left for you, see exercise **12.11**.

For future use, applying (12.9) at times t and T, we obtain that

$$X_T = X_t \exp\left((\alpha - \tfrac{1}{2}\sigma^2)(T-t) + \sigma(B_T - B_t)\right). \tag{12.10}$$

12.3 Stochastic exponentials and martingale representation

In this section we look at a close relative of the SDE (12.8) for geometric Brownian motion. In particular, we look at

$$dX_t = \sigma_t X_t \, dB_t \tag{12.11}$$

with the initial condition $X_0 = 1$. Here, σ_t is a stochastic process. By comparison to (12.8), we have set $\alpha = 0$ (which makes our life easier) but σ is no longer a deterministic constant (which makes our life harder).

The key idea is the same: we assume that a strictly positive solution exists, take logarithms $Z_t = \log X_t$, then look for an explicit formula for Z, and in turn an explicit formula for X, which we can then go back and check is really a solution.

From Ito's formula we have

$$\begin{aligned} dZ_t &= \left(0 + 0 + \frac{1}{2}(\sigma_t X_t)^2 \frac{-1}{X_t^2}\right) dt + \sigma_t X_t \frac{1}{X_t} \, dB_t \\ &= -\frac{1}{2}\sigma_t^2 \, dt + \sigma_t \, dB_t. \end{aligned}$$

This gives us

$$Z_t = Z_0 + \int_0^t \sigma_u \, dB_u - \frac{1}{2}\int_0^t \sigma_u^2 \, du$$

and hence

$$X_t = X_0 \exp\left(\int_0^t \sigma_u \, dB_u - \frac{1}{2}\int_0^t \sigma_u^2 \, du\right). \tag{12.12}$$

It can be checked (again, left for you, see **12.12**) that this formula really does solve (12.11).

Remark 12.3.1 If σ is a deterministic constant, then (12.12) becomes precisely (12.9) with $\alpha = 0$; as we would expect since this is also how (12.11) is connected too (12.8).

In view of (12.12) we have:

Definition 12.3.2 The *stochastic exponential* of the process σ_t is

$$\mathcal{E}_\sigma(t) = \exp\left(\int_0^t \sigma_u \, dB_u - \frac{1}{2}\int_0^t \sigma_u^2 \, du\right).$$

Of course, we have shown that $\mathcal{E}_\sigma(t)$ solves (12.11), and noting that $\mathcal{E}_\sigma(0) = 1$ we thus have

$$\mathcal{E}_\sigma(t) = 1 + \int_0^t \sigma_u \mathcal{E}_\sigma(u) \, dB_u. \tag{12.13}$$

We record this equation here because we'll need it in the next section.

The martingale representation theorem ($\star$)

Note that this section is off-syllabus, since it is marked with a ($\star$). However, since it covers a result that we will need in our analysis of the Black-Scholes model, it will still be covered in lectures.

Recall from Theorem 11.2.1 that Ito integrals $\int_0^t F_u\,dB_u$ are martingales. This might make us wonder if, given a martingale $M_t \in \mathcal{H}^2$, whether it is possible to write M as

$$M_t = M_0 + \int_0^t h_u\,dB_u$$

for some stochastic process h. Here, we follow common convention in denoting h_t with a lower case letter. The answer is strongly positive:

Theorem 12.3.3 (Martingale Representation Theorem) *Let $M_t \in \mathcal{H}^2$ be a continuous martingale. Fix $T \in (0,\infty)$. Then there exists a stochastic process $h_t \in \mathcal{H}^2$ such that*

$$M_t = M_0 + \int_0^t h_u\,dB_u$$

for all $t \in [0,T]$.

Sketch of Proof: Thanks to (12.13), we already know that this theorem holds if $M_0 = 1$ and M_t is the stochastic exponential of some stochastic process σ_t – in this case we take $h_t = \sigma_t \mathcal{E}_\sigma(t)$. The proof of the martingale representation theorem, which we don't include in this course, works by showing that *any* continuous martingale $M_t \in \mathcal{H}^2$ can be approximated by a sequence $M_t^{(n)}$ of continuous martingales that are themselves stochastic exponentials. As a consequence, the martingale representation theorem tells us that the process h_t exists, but does not provide us with a formula for h_t. ∎

The martingale representation theorem illustrates the importance of Ito integrals, and suggests that they are likely to be helpful in situations involving continuous time martingales. In fact, Theorem 12.3.3 will sit right at the heart of the argument that we will use (in Section 14.2) to show that hedging strategies exist in continuous time.

12.4 Exercises on Chapter 12

In all the following questions, B_t denotes a Brownian motion and $\mathcal{F}_t$ denotes its generated filtration.

On Ito's formula

12.1 Write the following equations in integral form.

(a) $dX_t = 2t\,dt + B_t\,dB_t$ over the time interval $[0,t]$,

(b) $dY_t = t\,dt$ over the time interval $[t,T]$.

Write down a differential equation satisfied by Y_t. Is X_t differentiable?

12.2 Apply Ito's formula to find an expression for the stochastic differential of $Z_t = t^3 X_t$, where $dX_t = \alpha\,dt + \beta\,dB_t$ and $\alpha, \beta \in \mathbb{R}$ are deterministic constants.

12.3 In each case, find the stochastic differential dZ_t, with coefficients in terms of t, B_t and Z_t.

(a) $Z_t = tB_t^2$

(b) $Z_t = e^{\alpha t}$, where $\alpha > 0$ is a deterministic constant.

(c) $Z_t = (X_t)^{-1}$, where $dX_t = t^2\,dt + B_t\,dB_t$.

(d) $Z_t = \sin(X_t)$, where $dX_t = \cos(X_t)\,dt + \cos(X_t)\,dB_t$.

12.4 Find dF_t where $F_t = B_t^n$, where $n \geq 2$. Hence, show that

$$\mathbb{E}[B_t^n] = \frac{n(n-1)}{2}\int_0^t \mathbb{E}[B_u^{n-2}]\,du.$$

Check that this is consistent with the formula obtained in part (c) of exercise 10.4.

12.5 Show that the following processes are martingales:

(a) $X_t = e^{t/2}\cos(B_t)$,

(b) $Y_t = (B_t + t)e^{-B_t - t/2}$.

12.6 Use Ito's formula to show that

$$tB_t = \int_0^t u\,dB_u + \int_0^t B_u\,du.$$

On stochastic differential equations

12.7 Suppose that X_t satisfies $X_0 = 1$ and $dX_t = (2+2t)\,dt + B_t\,dB_t$.

(a) Find $\mathbb{E}[X_t]$ as a function of t.

(b) Let $Y_t = X_t^2$. Calculate dY_t and hence find $\mathrm{var}(X_t)$ as a function of t.

(c) Suppose that X'_t satisfies $X'_t = 1$ and $dX'_t = (2+2t)\,dt + G_t\,dB_t$, where G_t is some unknown function of t and B_t. Based on your solutions to (a) and (b), comment on whether X_t and X'_t are likely to have the same mean and/or variance.

12.8 Suppose that X_t satisfies $X_0 = 1$ and $dX_t = \alpha X_t\,dt + \sigma_t\,dB_t$, where α is a deterministic constant and σ_t is a stochastic process. Find $\mathbb{E}[X_t]$ as a function of t.

12.9 Suppose that X_t satisfies $X_t = 1$ and $dX_t = X_t\,dB_t$. Show that $\mathrm{var}(X_t) = e^t - 1$.

12.10 Consider the stochastic differential equation

$$dX_t = 3X_t^{1/3}\,dt + 3X_t^{2/3}\,dB_t$$

with the initial condition $X_0 = 0$.

(a) Show that $Z_t = B_t^3$ is a solution of this equation.

(b) Can you think of another solution?

12.11 Check that (12.9) is a solution of (12.8).

12.12 Check that (12.12) is a solution of (12.11).

Challenge Questions

12.13 Fix $T > 0$.

(a) Let Y be an $\mathcal{F}_T$ measurable random variable such that $Y \in L^2$. Show that $M_t = \mathbb{E}[Y \mid \mathcal{F}_t]$ is a martingale for $t \in [0, T]$.

(b) You may assume that the stochastic process M_t in part (a) is continuous. Hence, for any given $Y \in \mathcal{F}_T$, the martingale representation theorem, from Section 12.3, tells us that there exists a stochastic process h_t such that

$$M_t = M_0 + \int_0^t h_u\,dB_u.$$

for all $t \in [0, T]$. Find an explicit formula for h_t in each of the following cases.

(i) $Y = B_T^2$

(ii) $Y = B_T^3$

(iii) $Y = e^{\sigma B_T}$, where $\sigma > 0$ is a deterministic constant.

Hint: Use the various connections that we've already found (in lemmas, exercises, examples, etc) between Brownian motion and martingales. For example, for (i) you might look at the formula for dZ_t where $Z_t = B_t^2$.

Chapter 13

Connections between SDEs and PDEs

We've already made several comparisons between ordinary differential equations and SDEs. We make a further connection in this section: we show how SDEs can be used to represent solutions to a particular family of partial differential equations.

13.1 The Feynman-Kac formula

Consider $F(t,x)$ where $t \in [0,T]$ and $x \in \mathbb{R}$. We will begin by looking at the partial differential equation

$$\frac{\partial F}{\partial t}(t,x) + \alpha(t,x)\frac{\partial F}{\partial x}(t,x) + \frac{1}{2}\beta(t,x)^2\frac{\partial^2 F}{\partial x^2}(t,x) = 0 \tag{13.1}$$

$$F(T,x) = \Phi(x) \tag{13.2}$$

Here, $\alpha(t,x)$ and $\beta(t,x)$ are (deterministic) continuous functions, and $\Phi(x)$ is a function known as the *boundary condition.*

Remark 13.1.1 ($\star$) Those familiar with partial differential equations may want to hear the 'proper' terminology: this is a second order parabolic PDE with a terminal boundary condition.

We will now show that the solutions of this PDE can be written in terms the solutions to a SDE. In particular, let X be a stochastic process that satisfies

$$dX_u = \alpha(u, X_u)\,dt + \beta(u, X_u)\,dB_u. \tag{13.3}$$

We will need to consider solutions to this SDE where we vary the initial value of x, and also the time at which the 'initial' value occurs. For given $x \in \mathbb{R}$ and $t \in [0,T]$ we write the subscripts $\mathbb{P}_{t,x}$ (and $\mathbb{E}_{t,x}$) to specify that X represents the solution of (13.3) with the initial condition that $X_t = x$ (and we are then interested in X_u during time $u \in [t,T]$). We will continue to use $\mathbb{P}$ and $\mathbb{E}$ to denote starting at time 0 with unspecified initial value X_0.

The connection is as follows.

Lemma 13.1.2 *Suppose that F is a solution of* (13.1) *and* (13.2)*, and also that $\beta(t,X_t)\frac{\partial F}{\partial x}(t,X_t)$ is in $\mathcal{H}^2$. Then,*

$$F(t,x) = \mathbb{E}_{t,x}\left[\Phi(X_T)\right]$$

for all $x \in \mathbb{R}$ and $t \in [0,T]$.

PROOF: We apply Ito's formula to $Z_t = F(t, X_t)$, giving

$$dZ_t = \left(\frac{\partial F}{\partial t} + \alpha(t, X_t)\frac{\partial F}{\partial x} + \frac{1}{2}\beta(t, X_t)^2\frac{\partial^2 F}{\partial x^2}\right) dt + \beta(t, X_t)\frac{\partial F}{\partial x} dB_t$$

where, as usual, we have suppressed the (t, X_t) arguments of the partial derivatives of F. We know that F satisfies the PDE (13.1), which means the first term on the right hand side of the above vanishes. Writing the result out with integrals, and taking the time limits to be $[t, T]$, then gives

$$F(T, X_T) = F(t, X_t) + \int_t^T \beta(u, X_u)\frac{\partial F}{\partial x}(u, X_u)\, dB_u.$$

We now take expectations $\mathbb{E}_{t,x}$, and recall from Theorem 11.2.1 that the expectation of integrals with respect to dB_t is zero. Note that here, to apply Theorem 11.2.1, we use that $\beta(t, X_t)\frac{\partial F}{\partial x}(u, X_u)$ is in $\mathcal{H}^2$. This leaves us with

$$\mathbb{E}_{t,x}[F(T, X_T)] = \mathbb{E}_{t,x}[F(t, X_t)]$$

Under $\mathbb{E}_{t,x}$ we have $X_t = x$, so $F(t, X_t) = F(t, x)$, which is deterministic. From (13.2) we have $F(T, X_T) = \Phi(T)$, which is also deterministic. Hence we have

$$\mathbb{E}_{t,x}[\Phi(X_T)] = F(t, x)$$

as required. ■

Lemma 13.1.2, (13.1) is very useful, from a theoretical point of view, but it is not quite what we need for later. The PDE that will turn out to be important for option pricing is

$$\frac{\partial F}{\partial t}(t, x) + \alpha(t, x)\frac{\partial F}{\partial x}(t, x) + \frac{1}{2}\beta(t, x)^2\frac{\partial^2 F}{\partial x^2}(t, x) - rF(t, x) = 0 \tag{13.4}$$

$$F(T, x) = \Phi(x) \tag{13.5}$$

where α, β, Φ are as before and r is a deterministic constant. We can treat this PDE in a similar style, even using the same SDE for X; we just need an extra term in the calculations.

Lemma 13.1.3 *Suppose that F is a solution of (13.4) and (13.5), and also that $\beta(t, X_t)\frac{\partial F}{\partial x}(t, X_t)$ is in $\mathcal{H}^2$. Then,*

$$F(t, x) = e^{-r(T-t)}\mathbb{E}_{t,x}\left[\Phi(X_T)\right] \tag{13.6}$$

for all $x \in \mathbb{R}$ and $t \in [0, T]$.

PROOF: This time we apply Ito's formula to the process $Z_t = e^{-rt}F(t, X_t)$. We obtain

$$\begin{aligned} dZ_t &= \left(-re^{-rt}F + e^{-rt}\frac{\partial F}{\partial t} + \alpha(t, X_t)e^{-rt}\frac{\partial F}{\partial x} + \frac{1}{2}\beta(t, X_t)^2 e^{-rt}\frac{\partial^2 F}{\partial x^2}\right) dt + \beta(t, X_t)e^{-rt}\frac{\partial F}{\partial x} dB_t \\ &= e^{-rt}\left(-rF + \frac{\partial F}{\partial t} + \alpha(t, X_t)\frac{\partial F}{\partial x} + \frac{1}{2}\beta(t, X_t)^2\frac{\partial^2 F}{\partial x^2}\right) dt + e^{-rt}\beta(t, X_t)\frac{\partial F}{\partial x} dB_t. \end{aligned}$$

We know that F satisfies (13.4), so the first term on the right hand side vanishes. Writing the result as integrals with time interval $[t, T]$ we obtain

$$e^{-rT}F(T, X_T) = e^{-rt}F(t, X_t) + \int_t^T e^{-ru}\beta(u, X_u)\frac{\partial F}{\partial x}(u, X_u)\, dB_u.$$

The second term on the right is an Ito integral, and hence has mean zero. Taking expectations $\mathbb{E}_{t,x}$ leaves us with

$$\mathbb{E}_{t,x}\left[e^{-rT}F(T,X_T)\right] = \mathbb{E}_{t,x}\left[e^{-rt}F(t,X_t)\right].$$

Under $\mathbb{E}_{t,x}$ we have $X_t = x$, so $F(t,X_t) = F(t,x)$ which is deterministic. We obtain that

$$e^{-r(T-t)}\mathbb{E}_{t,x}\left[F(T,X_T)\right] = F(t,x)$$

and using (13.5) then gives us

$$e^{-r(T-t)}\mathbb{E}_{t,x}\left[\Phi(X_T)\right] = F(t,x)$$

as required. ∎

Remark 13.1.4 Setting $r = 0$ in Lemma 13.1.3 gets us back to the statement of Lemma 13.1.2.

We can start to see the connection to finance emerging in (13.6). If we set $t = 0$, we obtain

$$F(0,x) = e^{-rT}\mathbb{E}_{0,x}[\Phi(X_T)]$$

which bears a resemblance to the risk-neutral valuation formula we found in Chapter 5. The connection will be explored when we come to apply Lemma 13.1.3, in Section 14.3.

Both Lemma 13.1.2 and 13.1.3 assert that, for some particular PDE, if a solution exists then it has a particular form. They assert uniqueness of solutions, without proving that a solution exists. In fact, in both cases, a (unique) solution does exist; this can be proved either using theory from the PDE world, or by using delicate real analysis to show explicitly that $\mathbb{E}_{t,x}[\Phi(X_T)]$ is differentiable. We don't include this step in our course.

Lemmas 13.1.2, 13.1.3 and variations on the same theme are collectively known as 'the' *Feynman-Kac formula*. They are named after the Richard Feynman (a theoretical physicist, famous for his work in quantum mechanics) and Mark Kac (pronounced "Kats", a mathematician famous for his contributions to probability theory).

Remark 13.1.5 ($\star$) There are many other families of PDEs that have relationships to other families of stochastic processes. These connections are exploited by researchers to transfer information and results between the 'dual' worlds of stochastic processes and PDEs.

Example 13.1.6 In cases where we can solve (13.3), we can use the Feynman-Kac formula to find explicit solutions to PDEs. For example, consider

$$\begin{aligned}\frac{\partial F}{\partial t}(t,x) + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial x^2}(t,x) &= 0\\ f(T,x) &= x^2.\end{aligned}$$

Here, $\sigma \geq 0$ is a deterministic constant and $t, x \in \mathbb{R}$.

We have $\alpha(t,x) = 0$, $\beta(t,x) = \sigma$ (both constants), and $\Phi(x) = x^2$. From Lemma 13.1.2 the solution is given by

$$F(t,x) = \mathbb{E}_{t,x}[X_T^2]$$

where $dX_u = 0\,du + \sigma\,dB_u = \sigma\,dB_u$. This means that

$$X_T = X_t + \sigma\int_t^T dB_u = X_t + \sigma(B_T - B_t).$$

Therefore,

$$
\begin{aligned}
F(t,x) &= \mathbb{E}_{t,x}\left[(X_t + \sigma(B_T - B_t))^2\right] \\
&= \mathbb{E}\left[(x + \sigma(B_T - B_t))^2\right] \\
&= \mathbb{E}\left[x^2 + \sigma^2(B_T - B_t)^2 + 2x\sigma(B_T - B_t)\right] \\
&= x^2 + \sigma^2(T - t).
\end{aligned}
$$

Here, we use that $X_t = x$ under $\mathbb{E}_{t,x}$ and that $B_T - B_t \sim B_{T-t} \sim N(0, T - t)$.

See exercises **13.1** and **13.2** for further examples of this method.

13.2 The Markov property

The notation $\mathbb{E}_{t,x}[\ldots]$ from Section 13.1, along with conditional expectation, allows us to formally express one of the most useful concepts in probability theory: the Markov property.

The idea of the Markov property is the following. Suppose that we have a stochastic process (F_t), and that we have waited up until time t, so the information visible to us is given by $\mathcal{F}_t$. We want to make a best guesses for some information about X_T where $T > t$. In symbols this means that we have chosen some (deterministic) function Φ and we are interested in

$$\mathbb{E}[\Phi(X_T)\,|\,\mathcal{F}_t].$$

In principle, we have access to all the information in $\mathcal{F}_t$. However, in many cases it holds that

$$\mathbb{E}[\Phi(X_T)\,|\,\mathcal{F}_t] = \mathbb{E}_{t,X_t}[\Phi(X_T)]. \tag{13.7}$$

Recall our intuition for conditional expectation: we view the left hand side of (13.7) as our best guess for $\Phi(X_T)$, based on the information we have seen during $[0,t]$. On the right hand side, we simply start at time t, fix the value of F_t, run the stochastic process until time t, and take expectations. Thus, the right hand side relies on much less information (in particular, we ignore the values of X_u for $u \in [0,t)$ and only need to know X_t).

Equation (13.7) is known as the *Markov property* for the stochastic process X. Not all stochastic processes are Markov, but many of the most useful ones are. Intuitively, the future (random) behaviour of a Markov process depends on its current value – but, crucially, the future doesn't depend on the *whole* history of the process, just on the current value.

For our purposes we need only know that:

Lemma 13.2.1 *All Ito processes satisfy the Markov property.*

In particular, the formula (13.7) holds when X is Brownian motion, and when X is geometric Brownian motion.

13.3 Exercises on Chapter 13

On the Feymann-Kac formula

13.1 Find an explicit formula for the solution of the PDE

$$\frac{\partial f}{\partial t}(t,x) - 2t\,\frac{\partial f}{\partial x}(t,x) = 0$$
$$f(T,x) = e^x.$$

13.2 Find an explicit formula for the solution of the PDE

$$\frac{\partial f}{\partial t}(t,x) + \frac{\partial f}{\partial x}(t,x) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,x) = 0$$
$$f(T,x) = x^2.$$

13.3 Let $T>0$. Let F be the solution of the PDE

$$\frac{\partial F}{\partial t}(t,x) + \alpha(t,x)\frac{\partial F}{\partial x}(t,x) + \frac{1}{2}\beta(t,x)^2\frac{\partial^2 F}{\partial x^2}(t,x) + \frac{\partial \gamma}{\partial t}(t) = 0 \tag{13.8}$$
$$F(T,x) = \Phi(x). \tag{13.9}$$

Here, $\alpha(t,x)$, $\beta(t,x)$, $\gamma(t)$ and $\Phi(x)$ are known functions.

(a) Let X_t satisfy $dX_u = \alpha(u,x)\,du + \beta(u,x)\,dB_u$. Define $Z_t = F(t,X_t) + \gamma(t)$. Use Ito's formula to find dZ_t.

(b) Show that $F(t,x) = \mathbb{E}_{t,x}\left[\Phi(X_T)\right] + \gamma(T) - \gamma(t)$.

Challenge Questions

13.4 Give an example of a stochastic process that is adapted to the generated filtration $\mathcal{F}_t$ of a Brownian motion B_t, but which does not satisfy the Markov property.

Chapter 14

The Black-Scholes model

Our discussion of finance, in continuous time, will be centred around the **Black-Scholes model**. The Black-Scholes model is, in some sense, the continuous time version of the binomial model from Section 5.4. Moving into continuous time has one big advantage: we can make our stock price process more realistic.

14.1 The Black-Scholes market

The **Black-Scholes market** contains two assets, cash and stock. In analogy to our discrete time model, cash earns interest at a deterministic rate, whereas the value of stock fluctuates randomly. As in discrete time, the model has some parameters: r, μ and σ, all real valued deterministic constants.

Here is the model:

- The value of a unit of **stock** at time t is S_t, where S_t is a geometric Brownian motion (from Section 12.2) given by the SDE

$$dS_t = \mu S_t\,dt + \sigma S_t\,dB_t, \tag{14.1}$$

  with initial value S_0. Here, of course, B_t is a Brownian motion. From (12.9) we know that the (unique) solution of this SDE is $S_t = S_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma B_t}$.

- If we hold x units of **cash** at the start of a time interval of length t, its final value will be xe^{rt}. This is the definition of 'cash earning interest at continuous rate $r > 0$ for time t'.

  A neater way of representing it is that we think of cash as an asset whose value changes over time. That is, the value of a 'unit of cash' at time t is given by

$$dC_t = rC_t\,dt. \tag{14.2}$$

  with initial condition $C_0 = 1$, and (unique) solution $C_t = e^{rt}$.

It might be appealing to 'divide by dt' and write (14.2) as an ODE, in the form $\frac{dC_t}{dt} = rC_t$ (with solution $C_t = C_0e^{rt}$). Justifying this step rigorously requires an application of the fundamental theorem of calculus. Whilst it would be mathematically correct, it is not what we want; (14.2) is better because its form is more compatible with the SDE (14.1).

Remark 14.1.1 We write the initial value of the stock as S_0. We will use s as a variable. Note that this is different to our use of s in discrete time, in which we set $s = S_0$.

As usual, we work over a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ where the filtration $\mathcal{F}_t$ is generated by the Brownian motion B_t, that is

$$\mathcal{F}_t = \sigma(B_u\,;\, u \leq t).$$

Here, B_t is the same Brownian motion that drives the random stock price in (14.1). Within the Black-Scholes model, B_t is the only source of randomness that we'll need.

As in the binomial model, we assume that we can borrow both cash and stock, and hold real valued amounts in each case. Thus, our definition of a portfolio remains the same as before:

Definition 14.1.2 A **portfolio** is a pair $h = (x, y)$ where $x \in \mathbb{R}$ denotes an amount of cash and $y \in \mathbb{R}$ denotes a number of units of stock. The value of this portfolio at time t is

$$V_t^h = xC_t + yS_t.$$

However, our definition of a portfolio *strategy* needs to be upgraded. Previously, at each time $t \in \mathbb{N}$ we had a round of buying/selling, and then in between times $t \mapsto t + 1$ we had stock/cash changing in value. Now, both these process must occur together, continuously. Happily, we have already developed the theoretical framework to do this:

Definition 14.1.3 A **portfolio strategy** is a pair of continuous stochastic processes $(h_t) = (x_t, y_t)$ where both x_t and y_t are adapted to the filtration $\mathcal{F}_t$. The value of (h_t) at time t is

$$V_t^h = x_tC_t + y_tS_t.$$

Note that here we break our usual convention of writing random quantities in capitals and deterministic quantities in lower case. The amounts of cash x_t and stock y_t that we hold at a given time are stochastic processes. (Just like in discrete time.)

We have required that our portfolio strategies be continuous. This assumption is helpful from a mathematical point of view, because we have only developed Ito integration for continuous processes, but it is not entirely realistic. In reality, it is possible to buy/sell large amounts of stock in a single transaction. We'll put this issue to one side for now, but we will return to discuss discontinuities and related matters (such as transaction costs) in Sections 15.1 and 15.5.

Definition 14.1.4 A portfolio strategy (h_t) is said to be **self-financing** if

$$dV_t^h = x_t\,dC_t + y_t\,dS_t. \tag{14.3}$$

We will need a little thought to understand why this definition captures the concept of a portfolio being self-financing. To do so, we need to think of a stochastic differential dX_t as 'the change in X over a short time interval. That is, if we choose time limits $[t, t + \delta]$ where δ is small then $Y_t dX_t$ represents $\int_t^{t+\delta} Y_u\,dX_u \approx Y_t(X_{t+\delta} - X_t)$. Note that we use Y_t and not $Y_{t+\delta}$ here to match the definition of the Ito integral in (11.12). So, approximately, (14.3) means that

$$V_{t+\delta}^h - V_t^h = x_t(C_{t+\delta} - C_t) + y_t(S_{t+\delta} - S_t)$$

This means that x_t and y_t are chosen in such a way as the changes in C and S entirely explain the variation in the value V^h during the time interval $[t, t + \delta]$. In other words, we haven't injected any value in, nor taken any value out. Formally, imposing this condition for all t in a limit as $\delta \downarrow 0$ results in (14.3).

Finally, and exactly as before:

Definition 14.1.5 We say that a self-financing portfolio strategy (h_t) is an **arbitrage possibility** if

$$\begin{aligned} V_0^h &= 0 \\ \mathbb{P}[V_t^h \geq 0] &= 1 \\ \mathbb{P}[V_t^h > 0] &> 0 \end{aligned}$$

Definition 14.1.6 A **contingent claim** with date of exercise T is any random variable of the form $X = \Phi(S_T)$, where $\Phi : [0, \infty) \to \mathbb{R}$ is a deterministic function.

We say that a portfolio strategy $h_t = (x_t, y_t)$ **hedges** (or **replicates**) X if (h_t) is self-financing and $V_T^h = X$.

For now, we restrict ourselves to portfolios $h_t = (x_t, y_t)$ containing only cash and stock. Later on, in Chapter 15, we will also consider portfolios that include financial derivatives (such as call/put options).

14.2 Completeness

We'll begin our analysis of the Black-Scholes model by showing that, in the Black-Scholes market, we can replicate any contingent claim that has a well-defined expectation. This is, essentially, what is meant by completeness in the Black-Scholes model.

Remark 14.2.1 The Black-Scholes model uses Ito integration. Up to now, whenever we wrote an Ito integral $\int_0^t F_t \, dB_t$ we were careful to check that $F \in \mathcal{H}^2$. From now on, we won't go to the trouble of checking this condition (although it does always hold). We'll say slightly more about this issue in Section 15.5

In the binomial model we discovered the importance of the so-called risk-neutral world, $\mathbb{Q}$. We will see that the corresponding concept is equally important in continuous time.

Definition 14.2.2 The **risk-netural world** $\mathbb{Q}$ is the probability measure under which S_t evolves according to the SDE

$$dS_t = rS_t \, dt + \sigma S_t \, dB_t. \tag{14.4}$$

Here, B_t has the same distribution (i.e. Brownian motion) under $\mathbb{Q}$ as in the 'real' world $\mathbb{P}$.

The key point here is that, compared to the 'real' dynamics of S_t, given in (14.1), we have replaced μ with r.

Definition 14.2.3 As in Section 12.2, we refer to r as the **drift** and to σ as the **volatility**.

The next lemma, at first glance, appears rather odd. It gives an awkward looking condition under which we can replicate a contingent claim. The point, as we will see immediately after, is that it turns out we can always satisfy this condition.

For reasons that will become clear in Section 15.5, in this section we'll tend to write X (instead of our usual $\Phi(S_T)$) for contingent claims.

Lemma 14.2.4 *Let $X \in \mathcal{F}_T$ be a contingent claim in the Black-Scholes market with exercise time T. Suppose that the stochastic process*

$$M_t = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[X \,|\, \mathcal{F}_t\right]$$

exists and has the stochastic differential

$$dM_t = f_t \, dZ_t \tag{14.5}$$

where $Z_t = e^{-rt}S_t$, for some (continuous, adapated) stochastic process f_t. Then there exists a replicating portfolio strategy for X.

PROOF: We are looking for a portfolio strategy $h_t = (x_t, y_t)$ such that both

$$V_T^h = X \qquad\qquad , \tag{14.6}$$

$$dV_t^h = x_t dC_t + y_t dS_t. \tag{14.7}$$

Here, the first equation is 'replication' and the second is 'self-financing'. The portfolio strategy we will use is

$$x_t = \left(M_t - e^{-rt} f_t S_t\right)$$

$$y_t = f_t.$$

It is immediate that x_t and y_t are continuous and adapted, so $h_t = (x_t, y_t)$ is a portfolio strategy and we must check (14.6) and (14.7) hold. Recalling that $C_t = e^{rt}$, we note that

$$\begin{aligned} V_t^h &= x_t C_t + y_t S_t \\ &= e^{rt} M_t - f_t S_t + f_t S_t \\ &= e^{rt} M_t. \end{aligned} \tag{14.8}$$

Hence $V_T^h = e^{rT} M_T = e^{rT} e^{-rT} \mathbb{E}^{\mathbb{Q}}[X \,|\, \mathcal{F}_T] = X$, because $X \in \mathcal{F}_T$. This checks (14.6) i.e. (h_t) replicates X.

Now, for (14.7). We need to apply Ito's formula to obtain dV_t^h, and we'll do it using (14.8). This means that we need to calculate dM_t, which we can do using (14.5), meaning that we'll need to start by finding an expression for dZ_t.

Using Ito's formula on $Z_t = e^{rt} S_t$ we obtain that

$$\begin{aligned} dZ_t &= \left(-rS_t e^{-rt} + \mu S_t e^{-rt} + 0\right) dt + \sigma S_t e^{-rt}\, dB_t \\ &= e^{-rt} S_t (\mu - r)\, dt + \sigma e^{-rt} S_t\, dB_t. \end{aligned}$$

This represents Z as an Ito process. Hence, using (14.5),

$$dM_t = e^{-rt} f_t S_t (\mu - r)\, dt + \sigma e^{-rt} f_t S_t\, dB_t$$

and we are now ready to apply Ito's formula to $V_t^h = e^{rt} M_t$. We obtain

$$\begin{aligned} dV_t^h &= \left(re^{rt} M_t + e^{-rt} f_t S_t (\mu - r) e^{rt} + 0\right) dt + \sigma e^{-rt} f_t S_t e^{rt}\, dB_t \\ &= \left(re^{rt} M_t - r f_t S_t\right) dt + f_t \left[\mu S_t\, dt + \sigma S_t\, dB_t\right] \\ &= \left(M_t - e^{-rt} f_t S_t\right) \left[re^{rt}\, dt\right] + f_t \left[\mu S_t\, dt + \sigma S_t\, dB_t\right]. \\ &= \left(M_t - e^{-rt} f_t S_t\right) dC_t + f_t\, dS_t. \end{aligned}$$

The second and third lines are collecting terms, so that in the final line we can use the definitions of dC_t and dS_t from (14.2) and (14.1). Recalling the definitions of x_t and y_t we now have

$$dV_t^h = x_t\, dC_t + y_t\, dS_t$$

and, as required, we have checked (14.7) i.e. (h_t) is self-financing. ∎

Theorem 14.2.5 *Let $X \in \mathcal{F}_T$ be a contingent claim in the Black-Scholes market with exercise time T. Suppose that $\mathbb{E}^{\mathbb{Q}}[|X|] < \infty$. Then there exists a replicating portfolio strategy for X.*

PROOF: We define $M_t = e^{-rT} \mathbb{E}^{\mathbb{Q}}[X \,|\, \mathcal{F}_t]$ and look to apply Lemma 14.2.4. Note that since $\mathbb{E}^{\mathbb{Q}}[|X|] < \infty$ this conditional expectation is well defined. From Example 3.3.9 (which is easily adapted to continuous time) we have that $\mathbb{E}[X \,|\, \mathcal{F}_t]$ is a martingale. Hence, since e^{-rT} is just a constant, we know that M_t is a martingale.

The key step is the next one: use the martingale representation theorem (Theorem 12.3.3), in the risk neutral world $\mathbb{Q}$, to say that under $\mathbb{Q}$ there exists a process g_t such that

$$dM_t = g_t\, dB_t. \tag{14.9}$$

Let $Z_t = e^{-rt}S_t$, as in Lemma 14.2.4. Under $\mathbb{Q}$, the dynamics of S are that $dS_t = rS_t\,dt + \sigma S_t\,dB_t$, so in world $\mathbb{Q}$ when we apply Ito's formula to Z_t we obtain

$$\begin{aligned} dZ_t &= \left(-re^{-rt}S_t + rS_te^{-rt} + 0\right)\,dt + \sigma e^{-rt}S_t\,dB_t \\ &= \sigma Z_t\,dB_t. \end{aligned}$$

Combining with (14.9) we obtain

$$dM_t = \frac{g_t}{\sigma Z_t}\sigma Z_t\,dB_t = \frac{g_t}{\sigma Z_t}\,dZ_t.$$

This shows that (14.5) holds, with $f_t = \frac{g_t}{\sigma Z_t}$. Finally, we apply Lemma 14.2.4 and deduce that there exists a replicating portfolio strategy for X. ∎

In some ways, Theorem 14.2.5 is very unsatisfactory. It asserts that (essentially, all) contingent claims can be replicated, but it doesn't tell us how to find a replicating portfolio. The root cause of this issue is that we used the martingale representation theorem – which told us the process g existed, but couldn't give us an explicit formula for g. Without calculating g, we can't calculate x_t, y_t either.

Of course, to trade in a real market, we would need a replicating portfolio $h_t = (x_t, y_t)$, or at the very least a way to numerically estimate one. With this in mind, in the next section, we will show how to find a replicating portfolio explicitly. The argument we will use to do so relies on *already knowing* that a replicating portfolio exists; for this reason, we needed to prove Theorem 14.2.5 first.

Another issue is that we have not addressed is to ask if our replicating portfolio is unique. Potentially, two different replicating portfolios could exist. We will show in the next section that the replicating portfolio is, in fact, unique.

14.3 The Black-Scholes equation

We now look at how a replicating portfolio can be found explicitly, for a given contingent claim $\Phi(S_T)$. We will, from now on, assume that our model is free of arbitrage.

In the case of the binomial model we found ourselves solving pairs of linear equations for each time-step; now, in continuous time, we will instead find ourselves solving a partial differential equation. This is natural – our linear equations told us how things changed across a single time-step, and PDEs can describe the changing state of a system in continuous time.

Let $\Phi(S_T)$ be a contingent claim in the Black-Scholes market (with parameters r, μ, σ), and suppose that $F(t,s)$ is a (suitably differentiable) function such that $F(t,S_t)$ denotes the value of the contingent claim $\Phi(S_T)$ at time $t \in [0,T]$. Then, as we will show in Theorem 14.3.1, for all $s > 0$ and $t \in [0,T]$ we have

$$\frac{\partial F}{\partial t}(t,s) + rs\frac{\partial F}{\partial s}(t,s) + \frac{1}{2}s^2\sigma^2\frac{\partial^2 F}{\partial s^2}(t,s) - rF(t,s) = 0, \tag{14.10}$$

$$F(T,s) = \Phi(s). \tag{14.11}$$

This is known as the **Black-Scholes equation**. It dates from a now famous research article by Fischer Black and Myron Scholes[1], published in 1973. The rigorous mathematical basis for the model was provided, also in 1973, by Robert C. Merton[2] In 1997 Merton and Scholes received the Nobel prize in economics, in recognition of their contributions (Black died in 1995, and Nobel prizes are only awarded to the living).

Theorem 14.3.1 *Let $\Phi(S_T)$ be a contingent claim such that $\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)] < \infty$. Then a replicating portfolio for $\Phi(S_T)$ is given by*

$$x_t = \frac{1}{rC_t}\left(\frac{\partial F}{\partial t}(t,S_t) + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 F}{\partial s^2}(t,S_t)\right)$$
$$y_t = \frac{\partial F}{\partial s}(t,S_t)$$

where $F(t,s)$ is the solution to (14.10),(14.11). *The value of this portfolio at time t is equal*

$$F(t,S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)\,|\,\mathcal{F}_t], \tag{14.12}$$

and in particular its value at time 0 *is*

$$e^{-rT}\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)]. \tag{14.13}$$

The formulae given for x_t, y_t are a big improvement on Theorem 14.2.5, because a computer can simulate solutions to (14.10), as well as their partial derivatives, without (much) difficulty. This allows computation of the replicating portfolio in real-time. You are not expected to memorize these formulae for $h_t = (x_t, y_t)$. The other formula in Theorem 14.3.1 can be found on the formula sheet, in Appendix E.

Of course, (14.13) is known as the **risk-neutral valuation formula**, in direct analogy to the discreet time version we found in Proposition 5.2.6.

[1] Black, Fischer; Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". Journal of Political Economy. 81 (3): 637–654.

[2] Merton, Robert C. (1973). "Theory of Rational Option Pricing". Bell Journal of Economics and Management Science. The RAND Corporation.

We will most of this section proving the claims in Theorem 14.3.1.

PROOF: First let us deduce that (14.11). Equation (14.11) is known as the boundary condition, because it relates to the exercise time T. It holds because, at the exercise time T, at which the value of S_T is known, to avoid arbitrage the value of the contingent claim $\Phi(S_T)$ must be equal to its price $F(T,S_T)$ i.e. $\Phi(S_T)=F(T,S_T)$. Since S_T may take any positive value we simply replace S_T with a general $s>0$.

Note that we don't need to worry about $s<0$ because S_t is a geometric Brownian motion, which (from Section 12.2) is always positive.

We now work towards proving that (14.10) holds. From Theorem 14.2.5 we know that $\Phi(S_T)$ can be replicated by a self-financing portfolio strategy $h_t=(x_t,y_t)$. Therefore, because we assume that our model is free of arbitrage, the value of this portfolio strategy at time t must be equal to $F(t,S_t)$:

$$F(t,S_t)=V^h_t=x_tC_t+y_tS_t \tag{14.14}$$

In particular, this implies that $dF(t,S_t)=dV^h_t$. We plan to calculate both these stochastic differentials, written out in full, and then use Lemma 11.4.5 to equate the dt and dB_t coefficients. It will turn out that this leads to precisely (14.10), and along the way we will discover formulae for x_t and y_t.

Since (h_t) is self-financing we have

$$dV^h_t=x_t\,dC_t+y_t\,dS_t.$$

Substituting in (14.1) and (14.2), this becomes

$$dV^h_t=(x_trC_t+y_t\mu S_t)\,dt+y_t\sigma S_t\,dB_t. \tag{14.15}$$

Next: recalling that $dS_t=\mu S_t\,dt+\sigma S_t\,dB_t$, by Ito's formula we have

$$dF(t,S_t)=\left(\frac{\partial F}{\partial t}+\mu S_t\frac{\partial F}{\partial s}+\frac{1}{2}\sigma^2S_t^2\frac{\partial^2F}{\partial s^2}\right)dt+\sigma S_t\frac{\partial F}{\partial s}\,dB_t \tag{14.16}$$

where we have suppressed the (t,S_t) arguments of the partial derivatives of F.

Remark 14.3.2 Here, we assume, without justifying ourselves, that $F(t,s)$ is differentiable (once in t and twice in s). This is a minor issue that can be dealt with using appropriate results from analysis, but it is beyond the scope of our course.

Equating the dB_t coefficients between (14.15) and (14.16) gives us that

$$y_t=\frac{\partial F}{\partial s}.$$

With this in hand, equating the dt coefficients gives us

$$\frac{\partial F}{\partial t}+\mu S_t\frac{\partial F}{\partial s}+\frac{1}{2}\sigma^2S_t^2\frac{\partial^2F}{\partial s^2}=x_trC_t+\frac{\partial F}{\partial s}\mu S_t \tag{14.17}$$

and the terms $\mu S_t\frac{\partial F}{\partial s}$ cancel giving

$$\frac{\partial F}{\partial t}+\frac{1}{2}\sigma^2S_t^2\frac{\partial^2F}{\partial s^2}=x_trC_t.$$

so as

$$x_t = \frac{\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial s^2}}{rC_t}.$$

Using (14.14) to substitute in for C_t in (14.17), we obtain

$$\begin{aligned}\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial s^2} &= x_t r\left(\frac{F - y_t S_t}{x_t}\right)\\ &= rF - rS_t\frac{\partial F}{\partial s}.\end{aligned}$$

This is equation (14.10), but with S_t in place of s. Equation (14.10) follows, because we can see from (12.9) that S_t takes values on the whole of the positive reals $(0,\infty)$. Note that we have also discovered formulae for x_t, y_t along the way, which must satisfy $F(t, S_t) = V_t^h$ because they were derived to satisfy (14.14).

Next we use Lemma 13.1.3, which tells us that the solution to be Black-Scholes equation can be written as

$$F(t,s) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,s}[\Phi(S_T)].$$

Note that here we take expectation in the risk neutral world $\mathbb{Q}$, in which S_t follows $dS_t = rS_t\,dt + \sigma S_t\,dB_t$ (in the notation of Lemma 13.1.3 we have $\alpha(t,s) = rs$ and $\beta(r,s) = \sigma s$). We deduce that the value at time t of the replicating portfolio is

$$F(t,S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,S_t}[\Phi(S_T)]$$

To finish the proof, we use that S_t is a Markov process. By the Markov property at time t, from Lemma 13.2.1, we have

$$\mathbb{E}^{\mathbb{Q}}_{t,S_t}[\Phi(S_T)] = \mathbb{E}^{\mathbb{Q}}[\Phi(S_T)\,|\,\mathcal{F}_t].$$

Setting $t = 0$, and recalling that $\mathcal{F}_0$ is the trivial σ-field (containing no information) we have

$$\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)\,|\,\mathcal{F}_0] = \mathbb{E}^{\mathbb{Q}}[\Phi(S_T)],$$

as required. ■

Corollary 14.3.3 *The replicating portfolio given in Theorem 14.3.1 for $\Phi(S_T)$ is unique.*

Proof: The calculations in the proof of Theorem 14.3.1 showed that any replicating portfolio (consisting solely of stocks and cash) could be written in terms of the solution to the Black-Scholes equation, as in the statement of Theorem 14.3.1. It is known from the world of PDEs that the Black-Scholes equation has a unique solution, so the replicating portfolio is also unique. ■

Remark 14.3.4 We have now seen that the Feynman-Kac formula is important to mathematical finance; it is the key to establishing the risk-neutral valuation formula (14.13). This formula links arbitrage free pricing theory to Brownian motion, through the stochastic differential equation $dS_t = \mu S_t\,dt + \sigma S_t\,dB_t$.

The Feynman-Kac formula has much wider applications too. We have already mentioned heat diffusion and movements of particles within fluids, and we could re-phrase our 'by hand' results from Section 10.3 in terms of the Feynman-Kac formula. We list two further examples here:

- It is used to describe solutions to the Schrödinger equation, a PDE which is the equivalent in quantum mechanics of Newton's second law (i.e. $F = ma$, the 'law of motion').
- It is used, in a variety of SDE based models, by mathematicians trying to model evolution, to analyse the positions and/or proportions of genes within a population.

There are many others – Brownian motion sits right at the heart of the physical world.

14.4 Martingales and 'the risk-neutral world'

In this section we give a (brief) explanation of where the term 'risk-neutral world' comes from. Offering explanations for mathematical wording is something of a dangerous game – in practice terminology often arises from accidents of history. There are many cases where "Someone's Theorem" was not discovered by the same Someone whose name is generally quoted.

As we saw in Proposition 5.5.6, in the risk neutral world $\mathbb{Q}$ the discounted stock price is a martingale. In continuous time we have the precise equivalent:

Lemma 14.4.1 *The stochastic process*

$$\frac{S_t}{C_t}$$

is a martingale in the risk-neutral world $\mathbb{Q}$.

PROOF: Recall that $C_t = e^{rt}$. Thus we are interested in the process $X_t = e^{-rt}S_t$. In the risk-neutral world $\mathbb{Q}$ the process S_t satisfies $dS_t = rS_t\,dt + \sigma S_t\,dB_t$, so by Ito's formula we have

$$\begin{aligned} dX_t &= \left(-re^{-rt}S_t + rS_te^{-rt} + 0\right)\,dt + \sigma S_t e^{-rt}\,dB_t \\ &= \sigma S_t e^{-rt}\,dB_t \end{aligned}$$

By Theorem 11.2.1, we have that X_t is a martingale (under $\mathbb{Q}$). ∎

In fact, more is true. Our next result says that, in the risk neutral world, once we have discounted for interest rates (i.e. divided by C_t), the price of *any* contingent claim is a martingale. As we saw in Section 3.3, martingales model fair games that are (on average) neither advantageous or disadvantageous to their players. The term 'risk-neutral' captures the fact that, inside the risk-neutral world, if we forget about interest rates, buying/selling on the stock market would be a fair game.

It is important to remember that we do not believe that the risk-neutral world is the real world. We believe in the absence of arbitrage, and we only care about the risk neutral world because it is useful when calculating arbitrage free prices.

Proposition 14.4.2 *Let* $\Phi(S_T)$ *be a contingent claim and let* Π_t *denote the price of this contingent claim at time* t. *Then*

$$\frac{\Pi_t}{C_t}$$

is a martingale in the risk-neutral world $\mathbb{Q}$.

PROOF: Our strategy, which is based on the proof of Lemma 14.4.1, is to set $X_t = \frac{\Pi_t}{C_t} = e^{-rt}\Pi_t$ and calculate dX_t using Itos formula. If X_t is to be a martingale, it should have the form $dX_t = (\ldots)\,dB_t$, in which case we can use Theorem 11.2.1 to finish the proof. We'll carry out the whole proof in the risk-neutral world $\mathbb{Q}$.

To avoid arbitrage, Π_t is equal to the value of the replicating portfolio for $\Phi(S_T)$ at time t. Hence, by Theorem 14.3.1,

$$\Pi_t = F(t, S_t). \tag{14.18}$$

Recall that, in the risk neutral world $\mathbb{Q}$, we have $dS_t = rS_t\,dt + \sigma S_t\,dB_t$. By Ito's formula (applied in the risk neutral world!), we have

$$d\Pi_t = \left(\frac{\partial F}{\partial t} + rS_t\frac{\partial F}{\partial x} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 F}{\partial x^2}\right)\,dt + \sigma S_t\frac{\partial F}{\partial x}\,dB_t$$

where we have suppressed the (t, S_t) arguments of F and its partial derivatives. We know that F satisfies the Black-Scholes PDE (14.10). Hence,

$$d\Pi_t = rF\,dt + \sigma S_t \frac{\partial F}{\partial x}\,dB_t$$

and using (14.18) again we have

$$d\Pi_t = r\Pi_t\,dt + \sigma S_t \frac{\partial F}{\partial x}\,dB_t$$

which represents Π_t as an Ito process.

We have $X_t = e^{-rt}\Pi_t$. Using Ito's formula again, this gives us

$$\begin{aligned} dX_t &= \left(-re^{-rt}\Pi_t + r\Pi_t e^{-rt} + 0\right)\,dt + \sigma S_t \frac{\partial F}{\partial x} e^{-rt}\,dB_t \\ &= \sigma S_t \frac{\partial F}{\partial x} e^{-rt}\,dB_t. \end{aligned}$$

Hence, by Theorem 11.2.1, X_t is a martingale (under $\mathbb{Q}$). ■

14.5 The Black-Scholes formula

In principle, Theorem 14.3.1 tells us the arbitrage free price, under the Black-Scholes model, of any contingent claim (such that $\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)]$ exists). In many cases the only way to evaluate (14.13) is with numerics. However, in some cases it is possible to derive an explicit formulae.

In this section, we find explicit formulae for the case of European call options. The result is often referred to as the **Black-Scholes formula**.

We begin from the formula given to us by Theorem 14.3.1: the price of the contingent claim $\Phi(S_T)$ at time 0 is

$$F(t,S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_T)\,|\,\mathcal{F}_t\right]. \tag{14.19}$$

Here, since we take expectation in the risk-neutral world $\mathbb{Q}$, the stock price process S_t has dynamics

$$dS_t = rS_t\,dt + \sigma S_t\,dB_t.$$

From (12.10) we have

$$\begin{aligned} S_T &= S_t \exp\Big((r-\tfrac{1}{2}\sigma^2)(T-t) + \sigma(B_T - B_t)\Big) && (14.20)\\ &= S_t e^Z && (14.21)\end{aligned}$$

where $Z \sim N\left[(r-\frac{1}{2}\sigma^2)(T-t), \sigma^2(T-t)\right] = N[u,v^2]$ is independent of $\mathcal{F}_t$. Here, we use u and v to keep our notation manageable.

Before we attempt to price a call option, let us work through a simpler case.

Example 14.5.1 We look to find the price, at time $t\in[0,T]$, of the contingent claim $\Phi(S_T) = 2S_T+1$. From (14.19), the price is

$$e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[2S_T+1\,|\,\mathcal{F}_t\right].$$

Using (14.21) we obtain

$$\begin{aligned} e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[2S_T+1\,|\,\mathcal{F}_t\right] &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[2S_te^Z+1\,|\,\mathcal{F}_t\right]\\ &= e^{-r(T-t)}\left(2S_t\mathbb{E}^{\mathbb{Q}}\left[e^Z\,|\,\mathcal{F}_t\right]+1\right)\\ &= e^{-r(T-t)}\left(2S_t\mathbb{E}^{\mathbb{Q}}\left[e^Z\right]+1\right)\\ &= e^{-r(T-t)}\left(2S_te^{(r-\frac{1}{2}\sigma^2)(T-t)+\frac{1}{2}\sigma^2(T-t)}+1\right)\\ &= e^{-r(T-t)}\left(2S_te^{r(T-t)}+1\right)\\ &= 2S_t + e^{-r(T-t)}\end{aligned}$$

Here, we use that S_t is $\mathcal{F}_t$ measurable, and that Z is independent of $\mathcal{F}_t$. We use (10.2) to calculate $\mathbb{E}[e^Z]$.

Note that the price obtained corresponds to the hedging strategy of, at time 0, owning two units of stock, plus e^{-rT} cash. The value of this portfolio at time t is then $2S_t+e^{rt}e^{-rT} = 2S_t+e^{-r(T-t)}$.

Using the (14.19) in combination with either (14.20) or (14.21) is usually the best way to compute explicit pricing formulae. See exercises **14.2-14.5** for more examples in the same style.

We'll give one more example in these notes. Consider the case of a call option with strike price K and exercise date T. This case is not easy, and it will involve some hefty calculations because of the max present in the contingent claim:

$$\Phi(S_T) = \max(S_T - K, 0).$$

We are now looking to evaluate

$$F(t, S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\max(S_T - K, 0) \,|\, \mathcal{F}_t].$$

Using (14.21), and writing $s = S_t$, gives us

$$\begin{aligned}
F(t,s) &= e^{-r(T-t)} \int_{-\infty}^{\infty} \Phi(se^z) f_Z(z)\, dz \\
&= e^{-r(T-t)} \left(0 + \int_{\log(K/s)}^{\infty} (se^z - K) f_Z(z)\, dz \right) \\
&= e^{-r(T-t)} \int_{\log(K/s)}^{\infty} (se^z - K) \frac{1}{\sqrt{2\pi v}} e^{-\frac{(z-u)^2}{2v^2}}\, dz \\
&= \overbrace{e^{-r(T-t)} \int_{\log(K/s)}^{\infty} se^z \frac{1}{\sqrt{2\pi v}} e^{-\frac{(z-u)^2}{2v^2}}\, dz}^{\mathcal{A}} - \overbrace{e^{-r(T-t)} \int_{\log(K/s)}^{\infty} K \frac{1}{\sqrt{2\pi v}} e^{-\frac{(z-u)^2}{2v^2}}\, dz}^{\mathcal{B}} \qquad (14.22)
\end{aligned}$$

Here, to deduce the second line, we split the integral into the cases $se^Z < K$ and $se^Z \geq K$. In the first case, $\Phi(se^Z) = \max(se^Z - K, 0) = 0$.

The strategy now is to treat $\mathcal{A}$ and $\mathcal{B}$ separately. We plan to re-write each of $\mathcal{A}$ and $\mathcal{B}$ in terms of $\mathcal{N}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2}\, dx$, the cumulative distribution function of the $N(0,1)$ distribution.

Let's look at the the $\mathcal{B}$ term first. Setting $y = -z$, we have

$$\begin{aligned}
\mathcal{B} &= \frac{K}{\sqrt{2\pi v}} e^{-r(T-t)} \int_{-\log(K/s)}^{-\infty} e^{-\frac{(-y-u)^2}{2v^2}} (-1)\, dy \\
&= \frac{K}{\sqrt{2\pi v}} e^{-r(T-t)} \int_{-\infty}^{\log(s/K)} e^{-\frac{(y+u)^2}{2v^2}}\, dy
\end{aligned}$$

and setting $x - \frac{y+u}{v}$ we have

$$\begin{aligned}
\mathcal{B} &= \frac{K}{\sqrt{2\pi v}} e^{-r(T-t)} \int_{-\infty}^{\frac{1}{v}(\log(s/K)+u)} e^{-\frac{x^2}{2}} v\, dx \\
&= Ke^{-r(T-t)} \int_{-\infty}^{\frac{1}{v}(\log(s/K)+u)} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
&= Ke^{-r(T-t)} \mathcal{N}\left[\frac{1}{v} \left\{ \log\left(\frac{s}{K}\right) + u \right\} \right].
\end{aligned}$$

For the $\mathcal{A}$ term, we have an extra factor e^z. To handle this factor we will have to complete the square inside the exponential term before we make the x substitution (i.e. the same technique as in exercise **10.5**). Again setting $y = -z$ we have

$$\begin{aligned}
\mathcal{A} &= \frac{s}{\sqrt{2\pi v}} e^{-r(T-t)} \int_{-\log(K/s)}^{-\infty} e^{-y} e^{-\frac{(-y-u)^2}{2v^2}} (-1)\, dy \\
&= \frac{s}{\sqrt{2\pi v}} e^{-r(T-t)} \int_{-\infty}^{\log(s/K)} e^{-y} e^{-\frac{(y+u)^2}{2v^2}}\, dy
\end{aligned}$$

$$
\begin{aligned}
&= \frac{s}{\sqrt{2\pi}v} e^{-r(T-t)} \int_{-\infty}^{\log(s/K)} \exp\left(\frac{-1}{2v^2}\left[y^2 + (2u+2v^2)y + u^2\right]\right) dy \\
&= \frac{s}{\sqrt{2\pi}v} e^{-r(T-t)} \int_{-\infty}^{\log(s/K)} \exp\left(\frac{-1}{2v^2}\left[\left(y + (u+v^2)\right)^2 - (u+v^2)^2 + u^2\right]\right) dy \\
&= \frac{s}{\sqrt{2\pi}v} e^{-r(T-t)+\frac{(u+v^2)^2-u^2}{2v^2}} \int_{-\infty}^{\log(s/K)} \exp\left(-\frac{(y+(u+v^2)^2)}{2v^2}\right) dy
\end{aligned}
$$

Noting that $r(T-t) = u + \frac{1}{2}v^2$, it is easily seen that $-r(T-t) + \frac{(u+v^2)^2-u^2}{2v^2} = 0$, so the first exponential term is simply equal to 1. Setting $x = \frac{y+(u+v^2)}{2v^2}$ we obtain

$$
\begin{aligned}
\mathcal{A} &= \frac{s}{\sqrt{2\pi}v} \int_{-\infty}^{\frac{1}{v}\left(\log(s/K)+u+v^2\right)} \exp\left(-\frac{x^2}{2}\right) v\,dx \\
&= s\int_{-\infty}^{\frac{1}{v}\left(\log(s/K)+u+v^2\right)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\
&= s\,\mathcal{N}\left[\frac{1}{v}\left\{\log\left(\frac{s}{K}\right) + u + v^2\right\}\right].
\end{aligned}
$$

Substituting back in for u and v, and recalling that we use the shorthand $s = S_t$, we obtain the following formula for price, at time $t \in [0,T]$, of a European call option with strike price K and exercise date T.

$$
F(t, S_t) = S_t\,\mathcal{N}[d_1] - Ke^{-r(T-t)}\,\mathcal{N}[d_2] \tag{14.23}
$$

where

$$
\begin{aligned}
d_1 &= \frac{1}{\sigma\sqrt{T-t}}\left\{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)\right\} \\
d_2 &= \frac{1}{\sigma\sqrt{T-t}}\left\{\log\left(\frac{S_t}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)\right\}.
\end{aligned}
$$

This expression for $F(t, S_t)$ is known as the Black-Scholes formula. It is historically significant, which is the only reason that we have included it within the course. The precise form of the formula is not very important, and you are certainly not expected to remember it. However, the discovery by Black and Scholes that the absence of arbitrage led to explicit pricing formulae for commonly used derivatives (such as calls and puts), revolutionized the financial world.

14.6 Exercises on Chapter 14

All questions refer to the Black-Scholes model, and use the the same notation as in the rest of this chapter.

On the Black Scholes model

14.1 Let $c \in \mathbb{R}$ be a deterministic constant. Show that the functions $f(t, s) = cs$ and $g(t, s) = ce^{rt}$ are both solutions of the Black-Scholes PDE (14.10).

14.2 Consider the contingent claim $\Phi(S_T) = K$. Use the risk neutral valuation formula (14.13) to show that its price, at time t, is $Ke^{-r(T-t)}$.

14.3 (a) Find the price, at time t, of the contingent claim $\Phi(S_T) = \log(S_T)$. Show that at time $t = 0$ this gives a price of

$$e^{-rT}\left(\log S_0 + (r - \tfrac{1}{2}\sigma^2)T\right).$$

(b) An excitable mathematician suggests the following hedging strategy:

> *"At time* 0*, I will buy* $\log S_0$ *units of stock. Then I'll wait until time* T*. The stock will be worth* S_T *then so then I will have* $\log(S_T)$ *worth in stock. This means that at time* 0 *the arbitrage free price of the contingent claim* $\log(S_T)$ *is actually* $\log S_0$*, and therefore the formula in part (a) can't be true."*

Where is the flaw in this argument?

14.4 Let $\beta \geq 2$.

(a) Calculate dY_t where $Y_t = S_t^\beta$, in the risk neutral world $\mathbb{Q}$. Hence, show that Y_t is a geometric Brownian motion and find its parameters.

(b) Show that the arbitrage free price, at time t, of the contingent claim $\Phi(S_T) = S_T^\beta$ is given by

$$S_t^\beta \exp\left\{-r(T-t)(1-\beta) - \tfrac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right\}.$$

14.5 Let $0 < \alpha < \beta$ and $K > 0$ be deterministic constants. A binary option is a contract with contingent claim

$$\Phi(S_T) = \begin{cases} K & \text{if } S_T \in [\alpha, \beta] \\ 0 & \text{otherwise.} \end{cases}$$

Find an explicit formula (in terms of the cumulative distribution function of the $N(0, 1)$ distribution) for the value of this contingent claim at time $t \in [0, T]$.

14.6 Find a deterministic process F_t such that $\frac{S_t}{F_t}$ is a martingale under $\mathbb{P}$.

14.7 Given a contingent claim $\Phi(S_T)$, let $\Pi_t(\Phi)$ denote the price of this contingent claim at time $t \in [0, T]$. Let $\Phi_1(S_T)$ and $\Phi_2(S_T)$ be two contingent claims, with exercise date T, and let $\alpha, \beta \in \mathbb{R}$ be deterministic constants. Show that

$$\Pi_t(\alpha\Phi_1 + \beta\Phi_2) = \alpha\Pi_t(\Phi_1) + \beta\Pi_t(\Phi_2)$$

for all $t \in [0, T]$.

Challenge Questions

14.8 Our excitable mathematician from exercise **14.3** is back. In response to **14.4**, they say

> *The situation in exercise 14.4, is equivalent to a 'new' Black-Scholes model in which the stock asset is not S_t, but Y_t. Since, by part (a) of 14.4, Y_t follows a geometric Brownian motion, this new model is actually just our usual Black-Scholes model but with different parameters. In this new model, we can hedge a single unit of the new 'Y_t' stock by simply buying a single unit of the new stock. So, the arbitrage free price of a single unit of the new stock will be $Y_t = S_t^\beta$ and the answer claimed in 14.4(b) is wrong.*

Where is the flaw in this argument?

Chapter 15

Application and extension of the Black-Scholes model

In this section we study the issue of how to make enough connections between the Black-Scholes model and reality that it can be used in the process of trading financial derivatives. We include some (perhaps surprising) information about what the Black-Scholes model is used for, in practice.

15.1 Transaction costs and parity relations

In our definition of the Black-Scholes market we assumed that it was possible to buy and sell, continuously, without any cost to doing so. In reality, there are costs incurred each time a stock is bought/sold, in the form of administrative cost and taxes.

For most contingent claims $\Phi(S_T)$, the replicating portfolio $h_t = (x_t, y_t)$ given by Theorem 14.3.1 changes continuously with time. This would mean continually incurring transaction costs, which is not desirable.

Our first idea for dealing with transaction costs comes form the (wishful) observation that it would be nice if we could replicate contingent claims with a constant portfolio $h_t = (x, y)$ that did not vary with time.

Definition 15.1.1 A constant portfolio is a portfolio that we buy at time 0, and which we then hold until time T.

That is, we don't trade stock for cash, or cash for stock, during $(0, T)$. Note that the hedging portfolios $h_t = (x_t, y_t)$ that we found in Theorem 14.3.1 are (typically) non-constant, since $y_t = \frac{\partial F}{\partial s}$ will generally not be constant. Worse, Corollary 14.3.3 showed that these hedging portfolios were unique: there are no other self-financing portfolio strategies, based only on cash and stock, which replicate $\Phi(S_T)$.

A possible way around this limitation is to allow ourselves to hold portfolios that include options, as well as just cash and stock. Let us illustrate this idea with European call/put options. First, we need some notation. Given a contingent claim $\Phi(S_T)$ with exercise date T we write $\Pi_t(\Phi)$ for the price of this contingent claim at time t, and $h^\Phi_t = (x^\Phi_t, y^\Phi_t)$ for its replicating portfolio.

In the case of European call/put options, the key to finding constant replicating portfolios is the 'put-call parity relation', which we'll now introduce (although we have touched on the discrete time version of it in exercise 5.4). Let Φ^{call} and Φ^{put} denote the contingent claims corresponding respectively to European call and put options, both with strike price K and exercise date T. Let Φ^{stock} and Φ^{cash} denote the contingent claims corresponding respectively to (the value at time T of) a single unit of stock, and a single unit of cash. Thus we have

$$\begin{aligned}\Phi^{cash}(S_T) &= 1\\ \Phi^{stock}(S_T) &= S_T\\ \Phi^{call}(S_T) &= \max(S_T - K, 0)\\ \Phi^{put}(S_T) &= \max(K - S_T, 0).\end{aligned}$$

The **put-call parity relation** states that

$$\Phi^{put}(S_T) = \Phi^{call}(S_T) + K\Phi^{cash}(S_T) - \Phi^{stock}(S_T). \tag{15.1}$$

This can be verified from the definitions of the functions, and considering the two cases $S_T \geq K$ and $S_T \leq K$ separately. Doing so is left for you, in exercise **15.1**. It follows (strictly speaking, using the result of exercise (**14.7**)) that if we buy a portfolio, at time 0, consisting of

- one European call option (with strike K and exercise T),
- Ke^{-rT} units of cash,
- minus one units of stock,

then at time T this portfolio will have the same payoff as a European put option (with strike K and exercise T).

We could rearrange (15.1) into the form $\Phi^{call}(S_T) = \dots$ and carry out the same procedure for a call option. So, we learn that if we allow ourselves to hold portfolios containing options, we can hedge European call/put options using constant portfolios. It's possible to hedge other types of contract too:

Example 15.1.2 Consider a contract with contingent claim

$$\Phi^{straddle}(S_T) = |S_T - K| = \begin{cases} K - S_T & \text{if } S_T \leq K \\ S_T - K & \text{if } S_T > K. \end{cases}$$

This type of contingent claim is known as a straddle, with strike price K and exercise time T. It is easy to see that the parity relation

$$\Phi^{straddle}(S_T) = \Phi^{put}(S_T) + \Phi^{call}(S_T)$$

holds. Hence, we can hedge a straddle by holding a portfolio of one call option, plus one put option, with the same strike price K and exercise date T.

See exercises **15.3**, **15.4** and **15.5** for more examples.

There is a drawback here. This hedging strategy requires that we purchase calls and puts (whenever we like), based on the stock S_t with a strike K and a exercise dates T of our own choosing. For calls/puts and equally common types of derivative, this is, broadly speaking, possible. For exotic types of option, even if we can find a suitable relation between payoffs in the style of (15.1) the derivative markets are often less fluid and the hedging portfolio we wish to buy may simply not be available for sale.

A 'next best' approach, for general contingent claims (of possibly exotic options), is to try and approximate a general contingent claim $\Phi(S_T)$ with a constant hedging portfolio consisting of cash, stock and a variety of call options with a variety of strike prices and exercise times. It turns out that this approximation is possible, at least in theory, with arbitrarily good precision for a large class of contingent claims. The formal argument, which we don't include in this course, relies on results from analysis concerning piecewise linear approximation of functions. Unfortunately, in most cases a large number/variety of call options are needed, and this greatly increases the transaction cost of just buying the portfolio at time 0.

In short, there is no easy answer here. Transaction costs have to be incurred at some point, and there is no 'automatic' strategy that is best used to minimize them.

15.2 The Greeks

As usual, let $F(t,s)$ be a differentiable function such that $F(t,S_t)$ denote the value, at time t, of a portfolio that replicates the contingent claim $\Phi(S_T)$. We adopt one key idea from the previous section: we allow this portfolio strategy to include options as well as cash and stock.

In this section we explore the sensitivity of replicating portfolios to changes in:

1. The price of the underlying stock S_t.
2. The model parameters r, μ and σ.

In the first case, we are interested to know, at a given time, how exposed our current portfolio is to changes in the asset price. That is, if the stock price were to quickly fall/rise, how much value are we likely to gain/lose?

In the second case, our concern is that the model parameters we're using may not be a good match for reality (or that reality may change so as new parameters are needed). This is a serious issue, since in practice the values used for r, μ, σ are obtained by statistical inference, and it is not easy process to obtain them.

Various derivatives of F are used to assess the sensitivity of the associated portfolio, and they are known collectively as **the Greeks**. They are

$$\begin{aligned}
\Delta &= \frac{\partial F}{\partial s} && (Delta)\\
\Gamma &= \frac{\partial^2 F}{\partial s^2} && (Gamma)\\
\Theta &= \frac{\partial F}{\partial t} && (Theta)\\
\rho &= \frac{\partial F}{\partial r} && (rho)\\
\mathcal{V} &= \frac{\partial F}{\partial \sigma} && (Vega)
\end{aligned}$$

all of which are evaluated at (t, S_t). Note that, for ρ and $\mathcal{V}$, we regard r and σ as variables (instead of constants) and differentiate with respect to them. There is no point in having a derivative with respect to μ because, as we have seen in Theorem 14.3.1, μ does not affect the risk-neutral world and consequently its value has no effect on arbitrage free prices: $\frac{\partial F}{\partial \mu} = 0$.

If we have an explicit formula for F, such as the Black-Scholes formula (14.23) for European call options, then we can differentiate to find explicit formulae for the Greeks. This can involve some quite messy calculations, so we don't focus on this aspect of the Greeks in this course (but, see exercise **15.9** if you like doing messy calculations). In general, they can be estimated numerically.

For us, Δ and Γ are most important. In the next section we study hedging strategies based on Δ and Γ.

Remark 15.2.1 In fact, $\mathcal{V}$ is not a letter of the Greek alphabet, it is a calligraphic Latin V, but the terminology 'the Greeks' is used anyway.

15.3 Delta and Gamma Hedging

As in the previous section, consider a portfolio consisting of both options, stock and cash, whose value at time t is given by $F(t, S_t)$. In this section we will make heavy use of the first two Greeks, $\Delta(t, S_t) = \frac{\partial F}{\partial s}(t, S_t)$ and $\Gamma(t, S_t) = \frac{\partial^2 F}{\partial s^2}(t, S_t)$.

In this section we will need to consider several portfolios at once. If $F(t, S_t)$ is the value of a portfolio at time t then we say that F is the *price function* of this portfolio, and we write the corresponding Δ and Γ as

$$\Delta_F = \frac{\partial F}{\partial s}, \qquad\qquad \Gamma_F = \frac{\partial^2 F}{\partial s^2}.$$

We focus first on Δ_F.

Definition 15.3.1 A portfolio with price function F is said to be **delta neutral** at time t if $\Delta_F(t, S_t) = 0$.

Let us think about delta neutrality for a moment. In words, $\Delta_F(t, S_t) = 0$ says that the derivative, with respect to the stock price, of the value of the replicating portfolio F, is zero. This means that if S_t where to vary (slightly), we would not expect the value of F to vary much. In other words, to some extent, the value of F is not exposed to changes in the price of the underlying stock; which is a good thing, since it means the holder of the portfolio takes less risk. With this motivation, we will now look at a hedging strategy that tries to keep $\Delta \approx 0$.

For a contingent claim $\Phi(S_T)$, let $F(t, S_t)$ be the value of the 'usual' hedging portfolio $h_t = (x_t, y_t)$, consisting of just cash and stocks, that is provided by Theorem 14.3.1. Such a portfolio is typically not delta neutral. We consider including an amount z_t of some derivative (say, a call option) with itself has hedging portfolio with value $Z(t, S_t)$. The value of our new portfolio is therefore

$$V(t, S_t) = F(t, S_t) + z_t Z(t, S_t).$$

We would like this new portfolio to be delta neutral, say at time t. That is, we would like $\Delta_V = \frac{\partial V}{\partial s} = 0$, which gives us the equation

$$\frac{\partial F}{\partial s} + z_t \frac{\partial Z}{\partial s} = 0 \tag{15.2}$$

and solving for z_t we see that we should hold

$$z_t = -\frac{\Delta_F}{\Delta_Z}$$

units of the derivative. Adding this amount of the deriative into our usual replicating portfolio, with the aim of having $\Delta_V = 0$, is known as a **delta hedge**.

Example 15.3.2 Suppose that we have sold an option that has price $P(t, S_t)$ at time t, and we want to delta hedge the sale. So, we have $F(t, S_t) = -P(t, S_t)$.

Suppose that the 'derivative' that we wish to use for our delta hedge is the underlying stock itself, which of course has price function

$$Z(t, S_t) = S_t,$$

giving $Z(t, s) = s$. Then $\frac{\partial V}{\partial s} = 0$ when

$$\frac{\partial F}{\partial s} + z_t \frac{\partial Z}{\partial s} = -\frac{\partial P}{\partial s} + z_t = 0$$

and we see that we need to hold an additional

$$z_t = \Delta_P$$

units of stock in order to delta hedge.

There is a spanner in the works here. The value of z_t changes with time. If, at time t, we delta hedge using z_t, then at time $t+\epsilon$ we will discover that $z_{t+\epsilon}$ is slightly different from z_t and our delta hedge is no longer working. On the other hand, if we continually adapt our portfolio to precisely match z_t then we will suffer high transaction costs.

To handle this issue, there is a procedure known as **discrete rebalanced delta hedge**. We explain it in the setting of Example 15.3.2, where we have sold one unit of an option with price function P, and wish to delta hedge the sale.

First, we fix some $\epsilon > 0$. Then:

- At time $t = 0$, sell one unit of an option with price P.
- Compute $z_t = \Delta_P(t, S_t)$ (using Theorem 14.3.1) at $t = 0$ and buy (or sell) this many units of stock.
- Wait for time ϵ. Recompute $z_t = \Delta_P(t, S_t)$ at time $t = \epsilon$, then buy/sell stock to re-balance the amount of 'extra' stock that we hold, to match this new amount.
- Repeat the rebalancing at each time $t = \epsilon, 2\epsilon, 3\epsilon, 4\epsilon, \ldots$ and so on.

Of course, a smaller ϵ results in closer approximation of $z_t \approx \Delta_P(t, S_t)$ and (consequently) a more effective delta hedge, but with higher transaction costs; a larger ϵ results in less effective delta hedge but lower transaction costs. This is natural – we can't expect to reduce risk for free.

Remark 15.3.3 ($\star$) It can be shown that as $\epsilon \to 0$, the resulting portfolio approximates the true delta hedged portfolio that corresponds, at all times, to holding z_t extra stock.

Of course, there is no need for all of our rebalancing time intervals to have length ϵ. In fact, if Δ_P is changing rapidly then we will need to rebalance frequently in order to keep $z_t \approx \Delta_P(t, S_t)$, but if Δ_P is relatively stable then we'll want to rebalance infrequently and spend less on transaction costs. This observation leads us on to the idea of Γ neutrality.

Definition 15.3.4 A portfolio with price function F is said to be **gamma neutral** at time t if $\Gamma_F(t, S_t) = 0$.

The key idea is that $\Gamma_P(t, S_t) = \frac{\partial}{\partial s}\Delta_P(t, S_t)$ measures how quickly Δ_P changes in response to changes in the underlying stock price S_t. When $\Gamma_P \approx 0$, we have that Δ_P does not change quickly in response to small changes in the stock price. For this reason, it is advantageous to hold portfolios which are delta neutral *and* gamma neutral. How to achieve this?

We now find ourselves wanting to augment a replicating portfolio (that, recall, has price function $F(t, S_t)$) into a portfolio with price function $V(t, S_t)$ in such a way as both

$$\Delta_V = \frac{\partial V}{\partial s} = 0, \qquad \Gamma_V = \frac{\partial^2 V}{\partial s^2} = 0. \tag{15.3}$$

It should be intuitively clear that, because we now have two conditions to satisfy, we'll need to consider adding in two extra quantities in order to achieve this. So, consider adding in w_t of some derivative with price function $W(t,S_t)$ and z_t of some other derivative with price function $Z(t,S_t)$. Then $V(t,S_t)=F(t,S_t)+w_tW(t,S_t)+z_tZ(t,S_t)$ and to satisfy (15.3) we need that both

$$\begin{aligned}\Delta_F + w_t\Delta_W + z_t\Delta_Z &= 0\\ \Gamma_F + w_t\Gamma_W + z_t\Gamma_Z &= 0.\end{aligned} \tag{15.4}$$

We could solve this pair of linear equations to find formulae for w_t and z_t, in terms of the Δs and Γs. Since the formulae themselves are not particularly interesting to see, we won't bother. Including the resulting amounts w_t of W, plus z_t of Z, into a portfolio, in order to achieve (15.3) is known as a **gamma hedge**.

Remark 15.3.5 Of course, z_t and w_t vary with time, which means that implementing a gamma hedge requires a discrete rebalancing scheme, in the same spirit as we described for the delta hedge. In the interests of brevity, we don't go into any further details on this point.

Delta and gamma hedging are the basis for many of the hedging strategies that are employed by investment banks and hedge funds.

15.4 Exercises on Chapter 15

On parity relations

15.1 (a) Draw graphs of the functions $\Phi^{cash}(S_T)$, $\Phi^{stock}(S_T)$, $\Phi^{call}(S_T)$ and $\Phi^{put}(S_T)$ defined in Section 15.1, as functions of S_T.

(b) Verify that the put-call parity relation (15.1) holds.

15.2 Use the put-call parity relation to show that the price at time 0 of a European put option with strike price K and exercise date T is given by $Ke^{-rT}\mathcal{N}(-d_2) - S_0\mathcal{N}(-d_1)$, where d_1 and d_2 are from (14.23).

15.3 Write down a constant portfolio, which may consist of cash, stock and/or European call and put options, that would replicate the contingent claim

$$\begin{cases} S_T - 1 & \text{if } S_T > 1, \\ 0 & \text{if } S_T \in [-1, 1] \\ -1 - S_T & \text{if } S_T < -1. \end{cases}$$

15.4 Let $A \geq 0$ and $K \geq 0$ be deterministic constants. Consider the contingent claim, with exercise date T,

$$\Phi(S_T) = \begin{cases} K & \text{if } S_T \leq A, \\ K + A - S_T & \text{if } A \leq S_T \leq K + A, \\ 0 & \text{if } K + A < S_T. \end{cases}$$

(a) Sketch (for general $A \geq 0$) a graph of $\Phi(S_T)$ as a function of S_T. Find a constant portfolio consisting of European put options, with exercise dates and strike prices of your choice, that replicates $\Phi(S_T)$.

(b) Find a constant portfolio consisting of cash, stock, and European call options (with exercise dates and strike prices of your choice) that replicates $\Phi(S_T)$.

(c) In parts (a) and (b) we found two different replicating portfolios for $\Phi(S_T)$. However Corollary 14.3.3 claimed that 'replicating portfolios are unique'. Why is this not a contradiction?

15.5 Let A and B be deterministic constants with $A < B$. Consider the contingent claim

$$\Phi^{bull}(S_T) = \begin{cases} B & \text{if } S_T > B, \\ S_T & \text{if } A \leq S_T \leq B, \\ A & \text{if } S_T < A. \end{cases}$$

This is known as a 'bull spread'. Find a constant portfolio consisting of cash and call options that replicates $\Phi^{bull}(S_T)$.

On the Greeks and delta/gamma hedging

15.6 Let $\beta > 2$. Find the values of all the Greeks, at time $t \in [0, T]$, for the derivative with contingent claim $\Phi(S_T) = S_T^{\beta}$. (*Hint: You will need part (b) of 14.4.*)

15.7 At time t you hold a portfolio h with value $F(t, S_t)$, for which (at time t) $\Delta_F = 2$ and $\Gamma_F = 3$.

(a) You want to make this portfolio delta neutral by adding a quantity of the underlying stock S_t. How much should you add? What is the cost of doing so?

(b) You want to make this portfolio both delta and gamma neutral, by adding a combination of the underlying stock S_t as well a second financial derivative with value $D(t, S_t)$, for which $\Delta_D = 1$ and $\Gamma_D = 2$. How much of each should you add? What is the cost of doing so?

15.8 Consider trying to gamma hedge a portfolio with value $F(t, S_t)$, by adding in an amount w_t of a financial derivative with value $W(t, S_t)$ and an amount z_t of a financial derivative with value $Z(t, S_t)$.

(a) An excitable mathematician suggests the following idea:

First, delta hedge using W: add in $w_t = -\frac{\Delta_F}{\Delta_W}$ of the first derivative to make our portfolio delta neutral. Then, add in a suitable amount z_t of Z to make the portfolio gamma neutral.

Why does this idea not work?

(b) Consider the case in which $Z(t, S_t) = S_t$. In this case, solve the equations (15.4) to find explicit expressions for w_t and z_t.

(c) Does the following idea work? Explain why, or why not.

First, add in an amount w_t of the first derivative to make the portfolio gamma neutral. Then, add in a suitable amount z_t of stock to make the portfolio delta neutral.

Challenge Questions

15.9 Use the Black-Scholes formula (14.23) to verify that, in the case of a European call option with strike price K and exercise date T, the Greeks are given by

$$\Delta = \mathcal{N}(d_1), \qquad \Gamma = \frac{\phi(d_1)}{s\delta\sqrt{T-t}},$$

$$\rho = K(T-t)e^{-r(T-t)}\mathcal{N}(d_2), \qquad \Theta = -\frac{s\phi(d_1)\sigma}{2\sqrt{T-t}} - rKe^{-r(T-t)}\mathcal{N}(d_2), \qquad \mathcal{V} = s\phi(d_1)\sqrt{T-t}.$$

Here, $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ is the p.d.f. of a $N(0,1)$ distribution, and $\mathcal{N}(x)$ is its c.d.f.

Use the put-call parity relation to find the values of the Greeks in the case of a European put option.

15.5 Further extensions $(\star)$

In this section we briefly survey a number of further ways in which the standard Black-Scholes model is often extended. Again, we follow our strategy of keeping things simple by studying each extension in isolation.

Note that the whole section of Section 15.5 is off-syllabus, and marked with $(\star)$s.

Time inhomogeneity $(\star)$

We've kept the model parameters r, μ and σ as fixed, deterministic constants throughout our analysis of the Black-Scholes model. In fact, there is no need to do so. It is common to allow μ and σ to depend on both t and S_t, written $\mu(t, S_t)$ and $\sigma(t, S_t)$.

The situation for r is similar. It is common to allow r to depend on t, but not on S_t, written simply $r(t)$. The for this choice is simply that reason interest rates are generally not thought to be dependent on stock prices. Moreover, interest rates tend to vary much more slowly than stock prices, and it is not unusual to assume that r is constant.

Allowing r, μ and σ to vary makes the problem of parameter inference much more difficult, but it is easily absorbed (without major changes) into the pricing theory that we've studied in this course. Essentially, the reason that no major changes occur is that our use of Ito's formula did not ever require us to differentiate r, μ or σ.

In the situation where $r = r(t)$, $\mu = \mu(t, S_t)$ and $\sigma = \sigma(t, S_t)$, it turns out that the risk neutral valuation formula for the contingent claim $\Phi(S_T)$ turns into

$$\Pi_t = e^{-\int_t^T r(u)\,du}\,\mathbb{E}^{\mathbb{Q}}_{t,S_t}\left[\Phi(S_T)\right],$$

where the risk neutral world $\mathbb{Q}$ has the dynamics of S as

$$dS_t = r(t)S_t\,dt + \sigma(t, S_t)S_t\,dB_t.$$

It is easily seen that this generalizes the version of risk-neutral valuation that we proved in Theorem 14.3.1: when r is constant we have $\int_t^T r\,du = r(T-t)$. Note, though, that in this case we don't have the explicit formula (14.20) for S_T in terms of S_t.

American and Exotic options $(\star)$

We have generally used European call (and sometimes put) options as our canonical examples of financial derivatives. Whilst European call and put options are traded on many stocks, the most common type of financial derivative is actually the *American* call/put option.

In all our previous work, we assumed that the exercise time T of an option was agreed in advance, and was deterministic. **American options** are options in which the holder can choose when to exercise the option. For example, an American call option with strike K and 'final' exercise date T, gives the holder the right to buy one unit of stock S_t for a (pre-agreed, deterministic) strike price K *at a time of their own choosing* during $[0, T]$.

The time τ at which the holder chooses to exercise their right to buy may depend on the current value of the stock price. As a result, the argument that we used to derive the prices (in Section 14.3) breaks down. The underlying cause is that they relate to a family of PDEs that are much more difficult to handle than the standard Black-Scholes PDEs – and for which Feynman-Kac formulas are not known to exist. In some special cases (including the case of American call options) explicit hedging strategies that are known, which allow explicit formulas for prices to be found, but in general numerical techniques are the only option.

Remark 15.5.1 (Δ) To be precise, what American options lead too, in place of the risk-neutral valuation formula, is a so-called 'optimal stopping problem': their value is given by

$$\max_{\tau} \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}\Phi(S_\tau)\right] \tag{15.5}$$

where the max is taken over stopping times τ. The optimal stopping problem is to identify the stopping time τ at which the maximum occurs.

Optimal stopping problems are generally quite difficult, and the mathematics needed to attack (15.5) is outside of the scope of what we can cover in this course.

Another interesting class of financial derivatives are **exotic options**. This is a general term used for options in which the contingent claim is either complicated to write down, or simply of an unusual form. They include

1. **digital options**, which give a fixed payoff if (and only if) the stock price is above a particular threshold;

2. **barrier options**, in which the exercise rights of the holder vary according to whether the stock price has crossed particular thresholds;

3. **Asian options**, whose payoff depends on the average price of the underlying asset during a particular time period;

and many other variants (such as cliquet, rainbow, lookback, chooser, Bermudean, ...). Often such options are not traded on stock exchanges because there is insufficient demand for their individual characteristics. As a result, their prices are agreed through direct discussions between the two (or more) parties involved in the contract. We don't attempt to make a catalogue of the theory of pricing such options. If you want to see some examples, there are plenty in Chapters 11-14 of the book 'The Mathematics of Financial Derivatives' by Wilmott, Howison and Dewynne.

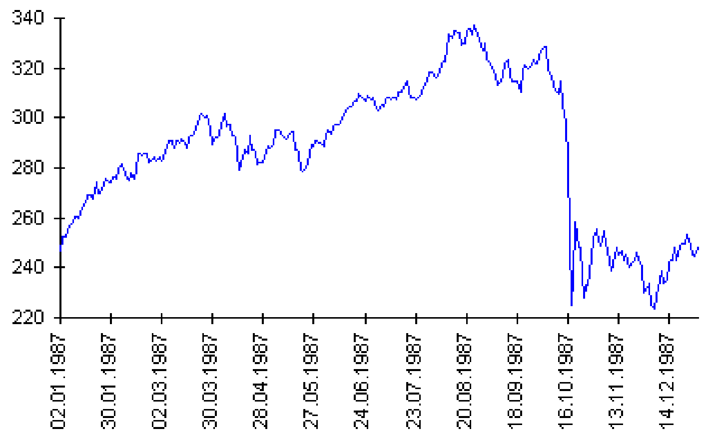
Remark 15.5.2 In Section 14.2 we were careful to write our contingent claims as X, rather than $\Phi(S_T)$. The reason is that the argument we gave for the Theorem 14.2.5, which essentially stated that the Black-Scholes market was complete, relied only on the fact that $X \in m\mathcal{F}_T$ (and not on having the form $X = \Phi(S_T)$).

As a consequence, we do know that most types of exotic option can be hedged in the Black-Scholes model – but, in general, we don't know how to find the replicating portfolios.

Discontinuous stock prices and heavy tails ($\star$)

We now begin to move further afield, towards some serious extensions of the standard Black-Scholes model.

The following graph shows the Standard & Poor's 500 index, usually known in short as the S&P 500, during 1987. It is essentially an averaged value of the stock prices of the top 500 companies within the American stock market[1]. The S&P 500 is widely regarded as one of the best ways of representing, in a single number, the value of stocks within the U.S. stock markets. We won't go into exactly how the 'average' is taken.



The feature of the graph that grabs our immediate attention is the huge fall, which occurs on Monday 19$^{\text{th}}$ October 1987. This date has become known as *Black Monday*. We see an instantaneous drop, with apparently no warning, in which the S&P 500 loses nearly 30% of its value. This is its largest fall ever, over twice the magnitude of the second largest fall (which you may remember: it occurred on 15$^{\text{th}}$ October 2008).

Remark 15.5.3 Something which may strike you as very strange: *no* single explanation for the Black Monday crash has ever been found. There are various theories as to what triggered the crash, ranging from a sudden lack liquidity to sudden implementations of new pricing methodology. It is often claimed that some *high-frequency trading algorithms* (which were relatively new, at the time) were programmed to automatically sell stocks when they saw them drop; meaning that once a small crash occurred there was suddenly a huge number of investors wanting to sell, exacerbating the drop in prices. However, in this case the largest drops occurred when trading volumes were low, meaning that high frequency trading is unlikely to be the full explanation.

Of the many stock market crashes that have occurred in history, one other deserves special mention: the financial crisis of 2007-8. We will discuss it (briefly) in Chapter 16. For now, let us focus on how we might incorporate rare, but sudden, downward drops in prices into the Black-Scholes model.

A rapid downward drop in prices is best represented by a discontinuity (downwards) in the stock price. However, the theory of Ito integration that we developed only worked for continuous stochastic process. In fact, a theory of Ito integration that can handle discontinuities does exist, but it requires much heavier use of analysis than the version we developed. The complication is that, in continuous time, we need to be clear about what information is known *instantaneously before* a jump takes place – we cannot allow ourselves to foresee the jump, since this would be unrealistic. Doing so necessitates much more careful use of σ-fields, filtrations and left/right-continuity than we saw in Chapter 11.

Happily, there is a generalization of Brownian motion, known as a *Lévy process*, which naturally incorporates unpredictable jumps, both upwards and/or downwards. It turns out that Lévy

[1]Including both the NYSE and the NASDAQ.

processes are intimately connected to heavy tailed random variables (that is, where $\mathbb{E}[X]$ or $\mathbb{E}[X^2]$ is not defined). This, in turn, means we need a further extension of Ito calculus, to remove our reliance on $\mathcal{H}^2$ and allow for infinite means and variances. After all this theoretical work, we can then build versions of the Black-Scholes model that incorporate the possibility of rare, but unpredictable, jumps in stock prices.

To summarize: such extensions are possible (and exist, and are used), but they are much harder to work with.

Volatility ($\star$)

Let us begin to think a little about what we would need to do to make use of the model, within a real market.

It is clear that we would need estimates of the parameters r and σ (we don't need to know μ, because it has no effect on arbitrage free prices). Estimating the interest rate r is often rather easy, because interests rates don't change quickly and they are *chosen* by banks, who generally also make them public information; so we won't worry about r.

It is much harder to estimate the volatility σ. We focus on this issue for the remainder of this section.

Historical volatility ($\star$)

One obvious idea is to estimate the future volatility based on the stock prices that we have observed in the recent past. Let us discuss one method for doing so.

Recall that our stock price process follows

$$dS_t = \mu S_t\,dt + \sigma S_t\,dB_t$$

which, as we saw in Section 12.2, has solution

$$S_t = S_0 \exp\left((\mu - \tfrac{1}{2}\sigma^2)t + \sigma B_t\right).$$

We fix $\epsilon > 0$ and set $t_i = i\epsilon$. We look at our historical data and find the values of S_t at times $t = t_i$, which we assume to all be in the past. Fix some $n \in \mathbb{N}$ and for $i = 1, \ldots, n$ define

$$\xi_i = \log\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right).$$

This gives us that

$$\xi_i = \left(\mu - \frac{\sigma^2}{2}\right)\epsilon + \sigma\left(B_{t_i} - B_{t_{i-1}}\right).$$

Using the independence properties of Brownian motion, we have that the $(\xi_i)_{i=1}^n$ are i.i.d. random variables with common distribution

$$N\left[(\mu - \tfrac{\sigma^2}{2})\epsilon,\ \sigma^2\epsilon\right].$$

We can estimate their (common) variance $\epsilon\sigma^2$ by the sample variance:

$$\epsilon\sigma^2 \approx \frac{1}{n-1}\sum_{i=1}^n \left(\xi_i - \bar{\xi}\right)^2$$

where $\bar{\xi} = \frac{1}{n}\sum_{i=1}^n \xi_i$.

Remark 15.5.4 If you are familiar with statistical inference, you should recognize the maximum likelihood estimator for the variance of the normal distribution.

We thus obtain the estimator

$$\hat{\sigma} = \frac{1}{\sqrt{\epsilon}} \left(\frac{1}{n-1} \sum_{i=1}^{n} (\xi_i - \bar{\xi})^2 \right)^{1/2}$$

for σ. What we have obtained here is a estimate of what the volatility was during the time period containing $t_1, \ldots, t_n$, which is in the past. When we are pricing options, what we would like to is an estimator of the volatility for future. We might be prepared to assume that the future will be similar to the past, and if we are then we can use $\hat{\sigma}$. Unfortunately, in most cases historical volatilities turn out to be a poor method of predicting future volatilities. With this in mind, we move on and examine an alternative idea.

Implied volatility and the volatility smile ($\star$)

Suppose that we wish to gauge what the (rest of the) market thinks is a reasonable estimate for volatility over the next, say, six months. Here's one way we could do it.

Take our Black-Scholes model. Find the pricing formula for a European call option that has a date of exercise six months into the future. Let us write this price as $c(K, t, T, r, \sigma)$. We know the value today of the stock S_t, we have argued that estimating r is not difficult, we know K, and we know that $T =$ six months. We can also *look at the market* and see the price at which this call option is being sold for – call this price p. We can then solve the equation

$$p = c(K, t, T, r, \sigma) \tag{15.6}$$

for σ, and obtain what is known as the **implied volatility**, often abbreviated simply to **vol**. Essentially, this is the volatility that 'the market' currently believes in.

This may seem like a circular procedure; if we were to then use the implied volatility as an estimator for σ and price accordingly, we would discover (at least, in theory) that we might as well have just charged whatever prices we could already see in the market. However, there is more to this situation that meets the eye.

Suppose that we are wanting to price an exotic derivative that is *not* commonly traded. This means that we cannot see current market prices for it – so we could not use our exotic derivative to find the implied volatility in the style of (15.6). But what we can do is:

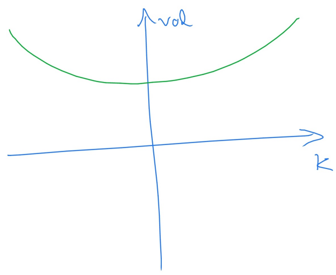
1. Look at prices of call options (or some other options, based on the same underlying stock, which are commonly traded) based on this derivative with the same exercise date. Use these to work out the implied volatility.

2. Use the resulting implied volatility, in the Black-Scholes model, to compute the price that we should charge for our exotic derivative.

In fact, *this* is currently the way in which the Black-Scholes model is most commonly used.

In another vein, we can use the idea of implied volatility to test the accuracy of the Black-Scholes model. Suppose that we observe the market prices of a number of call options, on the same stock, with the same exercise date, but with different strike prices. We can use each of these observations to calculate a value for the implied volatility. In theory, each of these calculations should give us the same σ.

They don't. In practice, it is often the case that options for which $S_t \gg K$ or $S_t \ll K$ (at the current time t) tend to suggest higher implied volatilities that those for which $S_t \approx K$. We thus obtain a graph of implied volatility as a function K, which typically looks like



This picture is known as the **volatility smile**. The exact shape of the smile varies from market to market, and can also vary substantially over time. Sometimes the smile becomes inverted, and is known as a 'volatility frown'.

The appearance of the volatility smile reflects deficiencies in the standard version of the Black-Scholes model. It is generally believed that the appearance of the volatility smile is closely connected to the standard Black-Scholes model not accounting for the possibility of jumps in the stock price; investigating this issue in detail is currently an active area of mathematical finance.

Despite the existence of the volatility smile, the Black-Scholes model is used extensively, in practice. The volatility smile provides an indicator of how well/badly the Black-Scholes model is capturing reality. It's consequences, including what can be learned from the precise shape of the smile, are well understood by traders.

It should be noted that, in practice, decisions taken on trading stocks and shares are based both on information obtained using modelling techniques (such as Black-Scholes) as well as qualitative information (such as reading the annual general reports of companies, being aware of the political environment, etc). Combining all this information together is a difficult task, which requires understanding the modelling theory well enough to judge, in detail, which aspects of reality are modelled well and which are not. There is nothing special to the world of finance here – all sophisticated stochastic modelling requires this the same level of care.

Incomplete markets ($\star$)

In Theorem 14.2.5 we proved that we could replicate any contingent claim X (providing that $\mathbb{E}^{\mathbb{Q}}[X]$ exists, which is not a major restriction). Essentially, replication relies on having enough independently tradeable commodities that we can create a portfolio that fully replicates the randomness built into the contingent claim X.

In some markets, the number of (independent) sources of random information is much greater than then number of (independent) commodities that are traded. They are known as **incomplete markets**. In practice, it is not particularly easy to judge if a given market fits into this class, but very many do.

In an incomplete market we can't hedge every contingent claim. Consequently, we also can't deduce arbitrage free prices for every contingent claim. Instead, it is possible to construct arbitrage based arguments to say that particular relationships exist *between* contingent claims must exist; statements of the form 'if the price of X is this then the price of Y must be that'. Then, by observing some carefully chosen prices from the market, we have enough information (when combined with the usual arbitrage free pricing methods) to uniquely determine the price of any contingent claim. The 'extra' information, that is observed via prices, essentially involves quantifying how risk averse the market is towards particular asset classes.

We won't go into any details on this procedure in these notes. A good source of further information is Chapter 15 of the book 'Arbitrage Theory in Continuous Time' by Bjork.

Chapter 16

The financial crisis of 2007/08 ($\star$)

In Section 15.5 we mentioned briefly that the cause of the Black Monday stock market crash, the single biggest fall in the S&P 500, has never been fully understood. The second, third, and fourth biggest one day falls in the S&P 500 all occurred towards the end of 2008, and form part of what is often referred to as the 'sub-prime mortgage crisis'. They present a very different picture. The U.S. Financial Crisis Enquiry Commission reported in 2011 that

> "the crisis was avoidable and was caused by: widespread failures in financial regulation, including the Federal Reserve's failure to stem the tide of toxic mortgages; dramatic breakdowns in corporate governance including too many financial firms acting recklessly and taking on too much risk; an explosive mix of excessive borrowing and risk by households and Wall Street that put the financial system on a collision course with crisis; key policy makers ill prepared for the crisis, lacking a full understanding of the financial system they oversaw; and systemic breaches in accountability and ethics at all levels."

Let us take some time to unpick the chain of events that occurred.

Availability of credit during 2000-07

During 2000-07 it was easy to obtain credit (i.e. borrow money) in Europe and the United States. Unrelated events in Russia and Asia during the late 1990s resulted in investors moving their money away, in many cases meaning they moved it into the U.S. and Europe. Part of this investment financed a boom in construction (i.e. house building) and also financed a boom in mortgages (i.e. loans with which people buy houses).

Loans have value: they are a contract which says that, at some point in the future, the money will be paid back, with interest. As a result loans are, essentially, a commodity. A 'share in a loan' is a share of the right to be re-paid when the time limit on the loan expires. These rights can be bought and sold.

As a consequence of these two factors, there was an increase in the number of financial derivatives for which the underlying 'stock' was mortgages. There was also a rise in house prices, because the easy availability of mortgages led to higher demand. We'll come back to this.

Deregulation of lending

In all economies, business that wish to grow require the ability to borrow money, in order to fund their growth. Investors (private investors, banks, governments, pension funds, etc.) provide this money, typically as a loan or by purchasing newly created stock. These investors take a risk: the value of their investment depends on the future success of the businesses in which they invest. Investment banks are one of the vehicles through which this process takes place. They act as middlemen, connecting multiple investors and business together.

Since the 1970s, governments in both the U.S. and Europe tended towards policies of *deregulation.* They aimed to offer more freedom to financial institutions, and (consequently) increase investment and activity within the wider economy.

Deregulation meant that financial institutions had to share less data about their own activities with regulators and policy makers. As a result, regulators did not immediately recognize the growth and increasing importance of investment banks and hedge funds to the wider economy. These institutions became major providers of credit, but were subject to less regulation than commercial banks. In part this lack of regulation was due to their use of complex financial derivatives, which were not well understood by regulators or subject to much regulation.

The housing boom

From around 2000 onwards, relaxed regulations allowed large numbers of lenders in U.S. to issue 'sub-prime' mortgages. These are mortgages issued to individuals who are at higher than normal risk of defaulting on their mortgage payments. When a homeowner is unable to keep up their mortgage payments, the bank takes ownership of their house (and its occupants must leave).

Of course, financial institutions are in a much better position than 'ordinary' people to predict, in the long run, whether or not someone is capable of paying back their mortgage – but (at least, initially) selling the mortgages was profitable, *including* selling the mortgages that were at high risk of eventually defaulting. As a result many financial institutions were keen to sell sub-prime mortgages. Moreover, deregulation allowed financial institutions to attract customers into 'variable rate' mortgages, that required lower repayments in their initial years, followed soon after by higher repayments.

At this point, you may hear alarm bells ringing and guess what happens next. Of course, you have the benefit of hindsight; in 2006/07 the prevalent view was that financial innovations were supporting a stable, high-growth housing market, with the (politically popular, and widely enjoyed) consequence of increased home-ownership.

The end of the housing boom

A house building boom, accompanied by a fast rise in house prices, does not continue forever. In around 2006, a point was reached where the supply of new houses outstripped demand for them, and house prices began to fall. Those who had taken out mortgages became less wealthy, since they owned a house whilst it decreased in value, but still owed the same mortgage repayments. In addition, the variable rate mortgages began to require higher repayments, with the result that many individuals (particularly in the U.S.) defaulted and lost their homes. This, in turn, increased the supply of empty houses, and further lowered house prices.

The mortgages (and the houses that became owned by banks when defaults occurred) were now worth much less. Some investment banks and other financial institutions simply ran out of money and collapsed. At this point, another unfortunate (and not foreseen) part of financial system came to light. Major investment banks had sold and re-sold large volumes of financial derivatives *to each other*, but there was no global register of who owed what to whom. Consequently, no-one knew which institution was most at risk of collapsing next, and no institution knew exactly which of its own investments were at risk of not being repaid. Worse, perhaps the collapse of a single large institution would result in enough unpaid debts that a chain of other institutions would be bought down with it.

The result was a situation where banks were very reluctant to lend, to anyone, including each other. This resulted in less investment in businesses, which as we have already commented, hurts the wider economy. Moreover, since lending is the major source of income for banks, a lack of lending results in all banks becoming weaker – and the cycle continues. Worse still, the divide between investment banks and commercial banks had gradually eroded: some of the institutions at risk of collapse were (also) high street banks, who hold the savings of the general public and operate ATMs. This situation, which occurred during 2008, seems good reason for the term 'the financial crisis'.

Outcomes

The possibility of the general public losing savings, along with wider effects on the economy, resulted in governments and central banks stepping in. Broadly speaking, they choose to provide loans and investment (using public money), to support financial institutions that were at risk of collapse and were also important to the public and the wider economy. Not all institutions were offered protection. We won't discuss the details of how these arrangements worked. In many countries, the amount of cash that was used to prop up failing financial institutions led to a very substantial deterioration in the state of public finances.

The impacts on the U.S. economy were huge. In the U.S. the S&P 500 lost around 45% of its value during 2008. The total value of houses within the U.S. dropped from \$13 trillion[1] in 2006 to less than \$9 trillion at the end of 2008. Total savings and investments owned by the general public, including retirement savings, dropped by around \$8.3 trillion (approximately $\$27,000$ *per person*, around 25% of an average persons savings).

Although the effects of the housing boom, and the housing boom itself, were greatest in the U.S., similar situations had occurred to varying extents in European countries. Coupled with the globally linked nature of the economy, and the fact that most major financial institutions now operate internationally, the financial crisis quickly spread to affect most of the developed world. A decrease in the amount of global trade resulted, along with a prolonged reduction in economic growth (which is still present today).

The process of changing the regulatory environment, in response to the financial crisis, is still ongoing. Broadly speaking, there is a move to involve *macroprudential* measures, which means considering the state of the financial system as a whole, instead of focusing in isolation only on the health of individual institutions. Mathematicians involved in this effort often argue that greater volumes of data on market activity should be collected and made public.

[1] $\$1$ trillion $= \$1,000,000,000,000$.

The role of financial derivatives

As we can see from this story, many different parties are involved and most of them can reasonably be attributed with a share of the blame. From our point of view, perhaps the most interesting aspect is that (with hindsight) it is clear that financial derivatives that were built on sub-prime mortgages were, prior to 2008, being priced incorrectly. These prices were typically computed by traders using variants and extensions of the Black-Scholes model.

The over-use of *collateralized debt obligations* (CDOs) is often cited as a practice that contributed greatly the financial crisis. CDOs are financial derivatives in which many different loans of varying quality are packaged together, and the holders of the CDO receive the repayments on the loans. Typically, not all loans are repaid and the CDO contract specifies that some of its holders are paid in preference to others.

Participation in a CDO is worth something and is therefore a tradeable asset. However, the underlying loans became packaged, bought, sold, divided, renamed, and repackaged, to such an extent that CDOs became highly complicated products in which the level of risk could not be known accurately. Moreover, packaging assets together in multi-party contracts increased the extent to which financial institutions became dependent on each other. Both these factors were not adequately captured by pricing models.

Another factor was *credit default swaps* (CDSs) in which an investor (typically an insurance company) would be paid an up-front sum, in cash, in return for promising to pay off the value of a loan in the case of a default. There was little regulation of this practice. Some institutions took on large volumes of mortgage CDSs and, in the short term, earned cash from doing so. Later, when mortgages defaulted, most were unable to cover their costs.

At the heart of the problem with CDSs was a widespread belief (before 2007) that the price of CDSs were correlated with the price of the underlying mortgages. This belief was realized through a pricing model known as the *Gaussian copula* formula, which became widely (and successfully) used during 2000-07. The intention of the model was to capture and predict correlations within the movements of prices of different assets. However, the shortcomings of the model were poorly understood and, in particular, it turned out to fail in the market conditions that emerged during 2007 when house prices dropped sharply.

Further reading

I recommend *The End of Alchemy: Money, Banking, and the Future of the Global Economy*, published in 2017 and written by Mervyn King, who was governor of the Bank of England from 2003 to 2013.

Chapter 17

Financial networks (Δ)

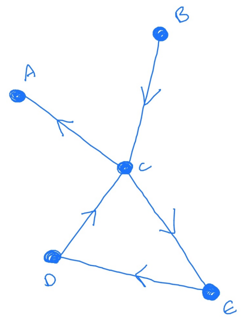
Following the financial crisis of 2007/08, both regulators and banks have become more interested in viewing the financial system as a connected whole (as opposed to viewing it as a collection of isolated institutions). With this new perspective, one aspect that commands special attention is *debt contagion*.

Debt contagion refers to the following scenario. Consider a connected network of banks who lend to each other. Suppose that one of these institutions, call it bank A, suddenly fails (i.e. goes bankrupt) and is then unable to pay its debts. In doing so, A harms its neighbours (i.e. banks who lent to A). Some of the neighbours of A may then also fail, and be unable to pay their own debts, harming their own neighbours – and so on. This process is usually known as a *cascade*. Potentially, the end result could be that a large fraction of the whole network fails.

As we mentioned in Chapter 16, it is thought that the financial system was at risk of precisely this scenario during parts of the financial crisis of 2008. Since then, there has been effort within the mathematical finance research community to provide models that describe when, and precisely how, such a risk is felt. In this section we give a brief introduction to one of the first (and consequently, simplest) models that was developed, followed by a discussion of how it was later extended.

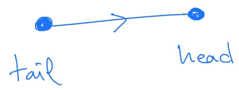
17.1 Graphs and random graphs (Δ)

A network is a set of nodes, some of which are connected together by edges. For example,

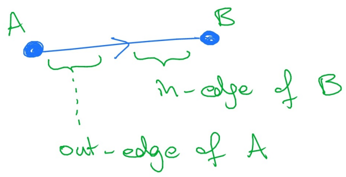


Here, the nodes are A, B, C, D, E. The terms *graph* and *network* are usually used interchangeably. The terms *vertex* and *node* are also used interchangeably. We'll use all these terms.

We will always be interested in the *directed* case, in which each edge has a direction. In the graph above this direction is signified by the direction of the arrows. The node at the start of an edge is called the head (of the edge) and the node at the end of the edge is called the tail (of the edge).



We write $\deg_{\text{in}}(i)$ for the number of edges that have node i as their head, and we call each such edge an *in-edge* of i. We write $\deg_{\text{out}}(i)$ for the number of edges that have i as their tail, and we call each such edge an *out-edge* of i. It's common to imagine each edge as split in half, with an 'in' part and an 'out' part



We write an edge as (tail, head). So the edge of the graph above are (C, A), (B, C), (C, E), (E, D) and (D, C).

Formally, a graph is a pair $G = (V, E)$ where V is the set of vertices and E is a set of ordered pairs of vertices. Each element of E has the form (v_1, v_2), where $v_1, v_2 \in V$, and denotes an edge with source v_1 and sink v_2. In this notation, our graph is

$$G = (V, E) \text{ where } V = \{A, B, C, D, E\} \text{ and } E = \{(C, A), (B, C), (C, E), (E, D), (D, C)\}.$$

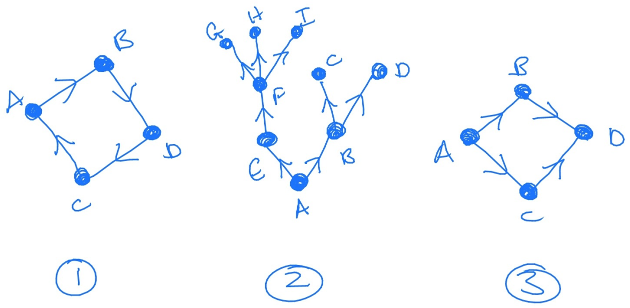
Given a graph $G = (V, E)$, the *degree distribution* of G is the distribution of the random variable

$$D_G = (\deg_{\text{in}}(v), \deg_{\text{out}}(v)), \tag{17.1}$$

where v is a node sampled uniformly at random from V. In words, D_G is the (bivariate) random variable who's distribution matches the frequencies of in/out degrees present in nodes of G. For example, for the graph above we have 5 nodes and

$$\begin{aligned}
\mathbb{P}[D_G = (1,0)] &= \frac{1}{5} && \text{(node } A\text{)}\\
\mathbb{P}[D_G = (0,1)] &= \frac{1}{5} && \text{(node } B\text{)}\\
\mathbb{P}[D_G = (2,2)] &= \frac{1}{5} && \text{(node } C\text{)}\\
\mathbb{P}[D_G = (1,1)] &= \frac{2}{5} && \text{(nodes } D \text{ and } E\text{)}
\end{aligned}$$

A *random graph* is, as you might expect, a graph where the sets of edges and vertices are both randomly sampled. We'll come back to thinking about random graphs in Section 17.3. A graph is said to be a *tree* if, between any pair A, B of vertices, there is precisely one path along edges (travelling in the direction they point) that gets from A to B. It is perhaps clearest from a picture:



Graph 1 is not a tree: for example, to get from A to A we can take the paths $ABDCA$ and $ABDCABDCA$. Graph 2 is a tree, because for any pair of vertices there is only one way to get between them. Graph 3 is not a tree: for example to get from A to D we can take the paths ABD and ACD.

17.2 The Gai-Kapadia model of debt contagion (Δ)

Fix $n \in \mathbb{N}$ and take a graph $G = (V, E)$. Think of each vertex $a \in V$ as a bank, and think of each edge $(a, b) \in E$ as saying that bank a has been loaned money by bank b.

Each loan has two possible states: healthy, or defaulted. Each bank has two possible states: healthy, or failed. Initially, all banks are assumed to be healthy, and all loans between all banks are assumed to be healthy.

We'll make the following key assumption, that there are numbers $\eta_j \in [0, 1]$ (known as contagion probabilities) such that:

(†) For any bank a, with in-degree j if, at any point, a is healthy and one of the loans owed to a becomes defaulted, then with probability η_j the bank a fails. All loans owed by bank a become defaulted.

Note that that banks who are owed a large number of loans (i.e. for which $\deg_{\text{in}}(a)$ is large) are less likely to fail when any *single* one of these loans becomes defaulted. We'll discuss how realistic this assumption is in Section 17.4.

The way the model works is as follows. To begin, a single bank is chosen uniformly at random; this bank fails and defaults on all of its loans. Because of (†) this (potentially) causes some other banks to default on their own loans, which in turn causes further defaults. We track the 'cascade' of defaults, as follows.

The set of loans which default at step $t = 0, 1, 2, \ldots,$ of the cascade will be denoted by L_t.

- On step $t = 0$, we pick a single bank a uniformly at random from $V = \{1, \ldots, n\}$. This banks fails, and all defaults on all of its loans: we set D_0 to be the set of out-edges of a.
- Then, iteratively, for $t = 1, 2, \ldots$, we construct L_t as follows.

  For each $(b, c) \in L_{t-1}$, we apply (†) to c. That is, if c is healthy then c fails with probability $\eta_{\deg_{\text{in}}(c)}$. If this causes c to fail, then we include all out-edges of c into L_t.

Eventually, because there are only finitely many edges in the graph, we reach a point at which L_t is empty. From then on, L_{t+1}, L_{t+2} are also empty. Then, the set of loans which were defaulted on during the cascade is

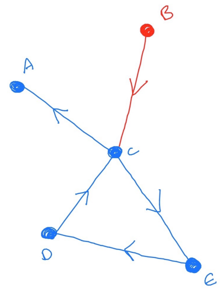
$$\mathscr{L} = \bigcup_{i=1}^{\infty} L_t. \tag{17.2}$$

We'll be interested in working out how big the set $\mathscr{L}$ is. That is, we are interested to know how many loans become defaulted once cascade finishes.

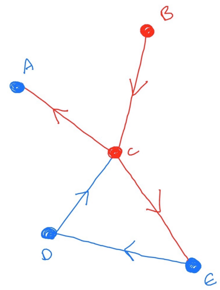
Remark 17.2.1 In reality, the main object of interest is how many banks fail during the cascade. This is closely connected to how many loans default but, for simplicity, we choose to focus on loans.

A summary of this model appears on the formula sheet, see Appendix E. We refer to it as the *Gai-Kapadia model (of debt contagion)*. The parameters η_j are known as the *contagion probabilities*.

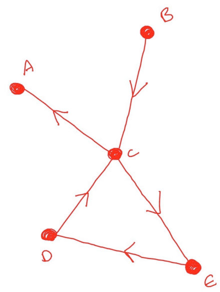
We'll now look at an example of the steps the cascade could take, on the graph we used as an example in Section 17.1. We'll mark defaulted/failed edges/nodes as red and take $\eta_j = \frac{1}{j}$. Initially, let us say that node B fails and defaults on the loan it owes to C:



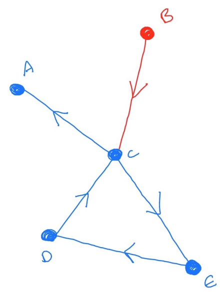
Since C has two in-edges, and $\eta_2 = \frac{1}{2}$, C has a probability $\frac{1}{2}$ of failing as a result of this default. So - we toss a coin and lets say that we discover that C does fail. Then C also defaults on all its loans:



Now, both E and A have only a single in-edge. Since $\eta_1 = 1$ this makes them certain to fail when their single debtor defaults. Similarly, D also fails. So the final result is that the whole graph is defaulted.



Of course, it didn't have to be this way. If our coin toss had gone the other way and C had not failed, the cascade would have finished after its first step and the final graph would simply look like



So we obtain that these two outcomes are both possible, each with probability $\frac{1}{2}$. Of course, in a bigger graph, the range of possible outcomes can be very large.

17.3 Approximating contagion by a Galton-Watson process (Δ)

Consider the debt contagion model from the previous section. We are interested to discover more about the quantity $\mathscr{L}$ defined in (17.2).

We need to specify which graph G we are using for the banking network. In fact, what we'll do is use some approximations. Firstly, we think of the cascade defined in Section 17.2 as exploring, edge by edge, a large *random* graph G. Recall the degree distribution D_G, defined in (17.1).

1. We'll imagine that our cascade explores a (large) random graph G. As we move through the graph, we map out the effects of the defaults.

In reality, to evaluate the cascade we would need to keep track of which nodes and edges we've already visited, but this makes our model too complicated. To make our calculations simpler:

2. We'll assume that each time we add a new out-edge into L_t, its associated in-edge is attached to a previously unseen node.

Each time we move along a defaulted edge into previously unseen node, we do not (by assumption!) encounter any of the nodes/edges previously involved in the cascade. So, following any given defaulted edge results in a random number G of new defaulted edges, independently of all else. This provides an important property: *the number of loans Z_n that are marked defaulted at each stage of the cascade is a Galton-Watson process.*

This assumption we've just made is an approximation – we are essentially approximating G with a randomly sampled tree. We need to be precise about how the sampling is done. At the same time, from Section 4.3 it is clear that we are most interested in knowing the expectation of the offspring distribution of the Galton-Watson process. In other words, we want to know, when we follow an edge of the cascade, the expected number of new edges that become added into the cascade.

Consider what would happen if we sampled an edge, uniformly at random from G, and moved along it. We use this as our approximation for what is found when we follow a defaulted edge. Where do we end up? Given a node $v \in G$, the number of in-edges of this node is $\deg_{\text{in}} v$. Sampling a random edge is equivalent to sampling a random in-edge, so the chance that we end up at a given node v is

$$\frac{\deg_{\text{in}}(v)}{\sum\limits_{u\in V}\deg_{\text{in}}(u)}.$$

The chance that this node fails as a result of our discovering it in our cascade (i.e. one of its in-edges defaults) is η_j, where $j = \deg_{in}(v)$. So the chance that we end up in v *and* that v is defaulted is

$$\frac{\deg_{\text{in}}(v)\eta_{\deg_{\text{in}}(v)}}{\sum\limits_{u\in V}\deg_{\text{in}}(u)}.$$

When this case occurs, all out-edges of v will become defaulted, which adds $\deg_{out}(v)$ new defaulted edges to our cascade. Hence, the expected number of newly defaulted edges that we discover in our cascade is

$$\sum_{v\in V}\deg_{\text{out}}(v)\frac{\deg_{\text{in}}(v)\eta_{\deg_{\text{in}}(v)}}{\sum\limits_{u\in V}\deg_{\text{in}}(u)}. \tag{17.3}$$

Let us write $|V|$ for the number of nodes in the graph and $|E|$ for the number of edges. Note that

$$|E| = \sum_{u \in V} \deg_{\text{in}}(u) = \sum_{u \in V} \deg_{\text{out}}(u).$$

We write $z = \frac{|E|}{|V|}$, for the average (in or out) degree of a vertex sampled uniformly at random from V.

Also, let us write $p_{j,k} = \mathbb{P}[D_G = (j,k)]$. By definition of D_G, the number of nodes with degree (j,k) is $|V|p_{j,k}$. Therefore, $\sum_{v \in V}(\ldots)$ is the same operation as $\sum_{j,k=0}^{\infty} |V|p_{j,k}(\ldots)$, where (j,k) represents the degree of node $v \in V$, and so we have

$$\begin{aligned}(17.3) &= \sum_{j,k=0}^{\infty} |V| p_{j,k} \frac{jk\eta_j}{|E|} \\ &= \frac{1}{z} \sum_{j,k=0}^{\infty} jk p_{j,k} \eta_j. \qquad (17.4)\end{aligned}$$

This is the expected number of newly defaulted loans that result from any given defaulted loans. If this quantity is strictly greater than one, then our Galton-Watson process has positive probability of tending to ∞ – meaning that our cascade of defaults can grow infinitely large. If not, then our Galton-Waton process only ever contains finitely many defaulted loans.

Therefore, our (approximate) analysis suggests that we could use the value of (17.4) as a criteria for how resilient our financial network is to debt contagion.

17.4 Modelling discussion on financial networks (Δ)

The analysis in Section 17.3 was essentially suggested in 2010 by Gai and Kapadia[1]. It was one of the earliest attempts at modelling debt contagion, and was published shortly after the financial crisis of 2008.

Typically, in the world of stochastic modelling, the first models of any new phenomenon are both simple and inaccurate; they provide a starting point for extension and refinement. With this in mind, let us discuss the shortcomings of the analysis in Section 17.3, and what might (and, in some cases, has) be done to improve it.

- The assumption (†) claims that a bank is equally dependent on each of its creditors. In practice, some loans are bigger than others, and some banks are better able to absorb defaults than others. For example, if it was the case that larger loans tended to be between larger (and, consequently, more strongly connected) banks, the model would not capture the effect.

  To correct this we'd want to understand the correlations between the size of a banks own balance sheet and the number of creditors/debtors it is connected too.

- The approximation used to turn the cascade into a Galton-Watson process results in us only ever visiting each bank once. This means that, in our approximation, each bank only ever sees *one* of its creditors default. As a result, we ignore the possibility that, once one of bank As creditors defaults it becomes very likely that other creditors of A will also default.

  For example, the real banking network could contain a core of strongly connected large banks all of whom lend large amounts to each other – and the effect of contagion essentially depends on how much it injures this central network. (In fact, in several cases where data on banking networks is available, this is now believed to be the case.)

One obvious question to ask is 'why not simply take the values from the real network and simulate a cascade of debt contagion on it?'. This does get done to some extent, with more complex models, but it is far from a complete answer to the problem. Regulators are interested in how you can modify the network (i.e. restructure the graph or change the rules) to become more resilient, and this demands a deeper understanding of the problem than simulations can typically provide. Also, simply running simulations and trusting them is very exposed to the shortcomings of the model, which itself may work well in some situations and badly in others. In practice, simulations and theoretical study are combined to provide insight into how resilient banking networks are.

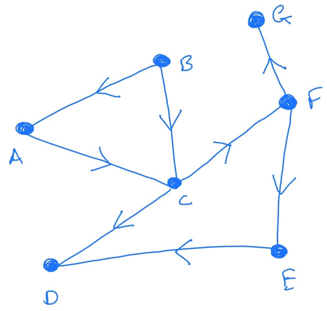
Similar (but not identical) methodology is used to model the spread of disease, the vulnerability of computer networks to hacking attacks, the propagation of news across social networks, and many other scenarios.

[1] Gai and Kapadia (2010), Contagion in financial networks, *Bank of England working paper*, Number 383.

17.5 Exercises on Chapter 17

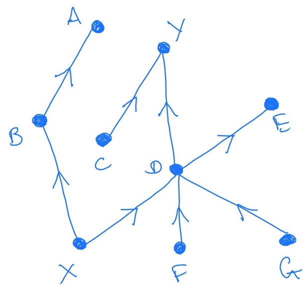
On random graphs and debt contagion

17.1 Consider the following graph G. Write down the distribution of D_G.



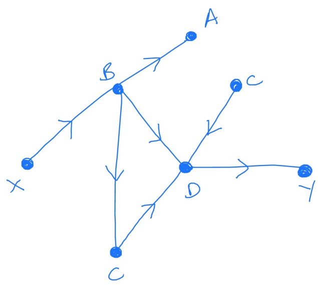
If we sample a uniformly random edge, and position ourselves at the head of this edge, what is the distribution of the out-degree O of the (random) node that we end up at?

17.2 Consider the following graph, as a banking network in the Gai-Kapadia model (as described in Section 17.2).



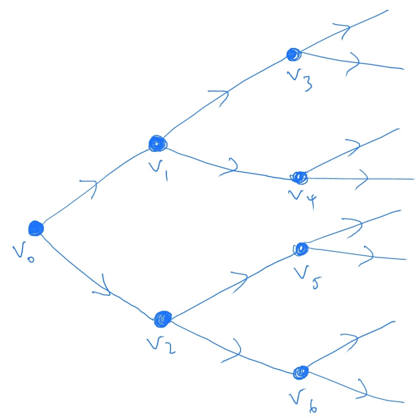
Let the contagion probabilities be $\eta_j = \frac{1}{j}$. Suppose that the bank marked X fails. What is the probability that the bank marked Y also fails?

17.3 Consider the following graph, as a banking network in the Gai-Kapadia model.



Let the contagion probabilities be $\eta_j = \frac{1}{j}$. Suppose that the bank marked X fails. What is the probability that the bank marked Y fails?

17.4 Consider the following graph (known as a binary tree) as a banking network in the Gai-Kapadia model.

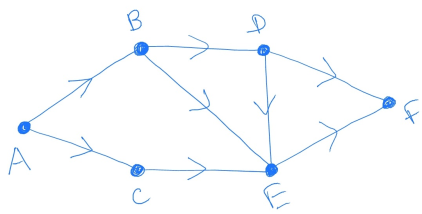


This graph is a tree, with infinitely many nodes, in which every node except for v_0 has one in-edge and two out-edges.

Let the contagion probabilities be $\eta_j = \alpha$, where $\alpha \in (0, 1)$ is constant. Suppose that the bank marked V_0 fails. Explain how the cascade of defaults that results can be represented as a Galton-Watson process.

We say that a 'catastrophic default' occurs if an infinite number of banks fail. Under what condition on α does this event have positive probability?

17.5 Consider the following graph, as a banking network in the Gai-Kapadia model.



Let the contagion probabilities be $\eta_j = \frac{1}{1+j}$. Suppose that the bank marked A fails.

(a) What is the probability that every bank within the graph fails?

(b) What is the probability that the bank marked F fails?

Appendix C

Solutions to exercises (book two)

Chapter 10

10.1 (a) We have $C_t = \mu t + \sigma B_t$. Hence,

$$\begin{aligned}\mathbb{E}[C_t] &= \mathbb{E}[\mu t + \sigma B_t] = \mu t + \sigma \mathbb{E}[B_t] = \mu t \\ \mathbb{E}[C_t^2] &= \mathbb{E}[\mu^2 t^t + 2t\mu\sigma B_t + \sigma^2 B_t^2] = \mu^2 t^2 + 2t\mu\sigma(0) + \sigma^2 t = \mu^2 t^2 + \sigma^2 t \\ \operatorname{var}(C_t) &= \mathbb{E}[C_t^2] - \mathbb{E}[C_t]^2 = \sigma^2 t.\end{aligned}$$

where we use that $\mathbb{E}[B_t] = 0$ and $\mathbb{E}[B_t^2] = t$.

(b) We have

$$C_t - C_u = \mu t + \sigma B_t - \mu u - \sigma B_u = \mu(t-u) + \sigma(B_t - B_u) \sim \mu(t-u) + \sigma N(0, t-u)$$

where, in the final step, we use the definition of Brownian motion. Then, by the scaling properties normal random variables we have $C_t - C_u \sim N(\mu(t-u), \sigma^2(t-u))$.

(c) Yes. By definition, Brownian motion B_t is a continuous stochastic process, meaning that the probability that B_t is a continuous function is one. Since $t \mapsto \mu t$ is a continuous function, we have that $\mu t + \sigma B_t$ is a continuous function with probability one; that is, B_t is a continuous stochastic process.

(d) We have $\mathbb{E}[C_t] = \mu t$, but Brownian motion has expectation zero, so C_t is not a Brownian motion.

10.2 We have

$$\begin{aligned}\operatorname{cov}(B_u, B_t) &= \mathbb{E}[B_t B_u] - \mathbb{E}[B_t]\mathbb{E}[B_u] \\ &= \mathbb{E}[B_t B_u] \\ &= \mathbb{E}[(B_t - B_u)B_u] + \mathbb{E}[B_u^2] \\ &= \mathbb{E}[B_t - B_u]\mathbb{E}[B_u] + \mathbb{E}[B_u^2] \\ &= 0 \times 0 + u \\ &= u\end{aligned}$$

Here, to deduce the fourth line, we use the second property in Theorem 10.2.1, which tells us that $B_t - B_u$ and B_u are independent.

10.3 When $u \leq t$ the martingale property of Brownian motion (Lemma 10.4.3) implies that $\mathbb{E}[B_t \,|\, \mathcal{F}_u] = B_u$. When $t \leq u$ we have $B_t \in \mathcal{F}_u$ so by taking out what is known we have $\mathbb{E}[B_t \,|\, \mathcal{F}_u] = B_t$. Combing the two cases, for all $u \geq 0$ and $t \geq 0$ we have $\mathbb{E}[B_t \,|\, \mathcal{F}_u] = B_{\min(u,t)}$.

10.4 (a) We'll use the pdf of the normal distribution to write $\mathbb{E}[B_t^n]$ as an integral. Then, integrating by parts (note that $\frac{d}{dz} e^{-\frac{z^2}{2t}} = \frac{-z}{t} e^{-\frac{z^2}{2t}}$) we have

$$\begin{aligned}\mathbb{E}[B_t^n] &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} z^n e^{-\frac{z^2}{2t}}\,dz \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \left(-tz^{n-1}\right)\left(\frac{-z}{t} e^{-\frac{z^2}{2t}}\right) dz\end{aligned}$$

$$
\begin{aligned}
&= \left[-\frac{1}{\sqrt{2\pi t}} t z^{n-1} e^{-\frac{z^2}{2t}}\right]_{z=-\infty}^{\infty} + \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} t(n-1) z^{n-2} t e^{-\frac{z^2}{2t}}\,dz \\
&= 0 + t(n-1)\mathbb{E}[B_t^{n-2}].
\end{aligned}
$$

(b) We use the formula we deduced in part (a). Since $\mathbb{E}[B_t^0] = \mathbb{E}[1] = 1$, we have $\mathbb{E}[B_t^2] = t(2-1)(1) = t$. Hence $\mathbb{E}[B_t^4] = t(4-1)\mathbb{E}[B_t^2] = t(4-1)t = 3t^2$ and therefore $\mathrm{var}(B_t^2) = \mathbb{E}[B_t^4] - \mathbb{E}[B_t^2]^2 = 2t^2$.

(c) Again, we use the formula we deduced in part (a). Since $\mathbb{E}[B_t] = 0$ it follows (by a trivial induction) that $\mathbb{E}[B_t^n] = 0$ for all odd $n \in \mathbb{N}$. For even $n \in \mathbb{N}$ we have $\mathbb{E}[B_t^2] = t$ and (by induction) we obtain

$$\mathbb{E}[B_t^n] = t^{n/2}(n-1)(n-3)\dots(1).$$

(d) We have $\mathrm{var}(B_t^n) = \mathbb{E}[B_t^{2n}] - \mathbb{E}[B_t^n]^2$ which is finite by part (c). Hence $B_t^n \in L^2$, which implies that $B_t^n \in L^1$.

10.5 Using the scaling properties of normal random variables, we write $Z = \mu + Y$ where $Y \sim N(0, \sigma^2)$. Then, $e^Z = e^\mu e^Y$ and

$$
\begin{aligned}
\mathbb{E}[e^Y] &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^y e^{-\frac{y^2}{2\sigma^2}}\,dy \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\left(y^2 - 2\sigma^2 y\right)\right\} dy \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\left((y-\sigma^2)^2 - \sigma^4\right)\right\} dy \\
&= e^{\frac{\sigma^2}{2}} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left\{-\frac{(y-\sigma^2)^2}{2\sigma^2}\right\} dy \\
&= e^{\frac{\sigma^2}{2}}.
\end{aligned}
$$

Here, to deduce the third line we complete the square, and to deduce the final line we use that the p.d.f. of a $N(\sigma^2, \sigma)$ random variable integrates to 1. Therefore,

$$\mathbb{E}[e^Z] = e^\mu \mathbb{E}[e^Y] = e^{\mu + \frac{\sigma^2}{2}}.$$

10.6 (a) Since B_t is adapted, $e^{\sigma B_t - \frac{1}{2}\sigma^2 t}$ is also adapted. By (10.2) and scaling of normal random variables we have that $e^{\sigma B_t - \frac{1}{2}\sigma^2 t}$ is in L^1. It remains only to check that

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right) \mid \mathcal{F}_u\right] &= \mathbb{E}\left[\exp\left(\sigma(B_t - B_u) + \sigma B_u - \frac{1}{2}\sigma^2 t\right) \mid \mathcal{F}_u\right] \\
&= \exp\left(\sigma B_u - \frac{1}{2}\sigma^2 t\right) \mathbb{E}\left[\exp\left(\sigma(B_t - B_u)\right) \mid \mathcal{F}_u\right] \\
&= \exp\left(\sigma B_u - \frac{1}{2}\sigma^2 t\right) \mathbb{E}\left[\exp\left(\sigma(B_t - B_u)\right)\right] \\
&= \exp\left(\sigma B_u - \frac{1}{2}\sigma^2 t\right) \exp\left(\frac{1}{2}\sigma^2(t-u)\right) \\
&= \exp\left(\sigma B_u - \frac{1}{2}\sigma^2 u\right).
\end{aligned}
$$

Here, the second line follows by taking out what is known, since B_u is $\mathcal{F}_u$ measurable. The third line then follows by the definition of Brownian motion, in particular that $B_t - B_u$ is independent of $\mathcal{F}_u$. The fourth line follows by (10.2), since $B_t - B_u \sim N(0, t-u)$ and hence $\sigma(B_t - B_u) \sim N(0, \sigma^2(t-u))$.

(b) Since B_t is adapted, $B_t^3 - 3tB_t$ is also adapted. From **10.4** we have $B_t^3, B_t \in L^1$, so also $B_t^3 - tB_t \in L^1$. Using that B_u is $\mathcal{F}_u$ measurable, we have

$$
\begin{aligned}
\mathbb{E}\left[B_t^3 - 3tB_t \,|\, \mathcal{F}_u\right] &= \mathbb{E}\left[(B_u^3 - 3uB_u) + B_t^3 - 3tB_t + (B_u^3 - 3uB_u) \,|\, \mathcal{F}_u\right] \\
&= B_u^3 - 3uB_u + \mathbb{E}\left[B_t^3 - B_u^3 - 3tB_t + 3uB_u \,|\, \mathcal{F}_u\right]
\end{aligned}
$$

so we need only check that the second term on the right hand side is zero. To see this,

$$
\begin{aligned}
\mathbb{E}\left[B_t^3 - B_u^3 - 3tB_t + 3uB_u \,|\, \mathcal{F}_u\right] &= \mathbb{E}\left[B_t^3 - B_u^3 - 3tB_u + 3uB_u - 3t(B_t - B_u) \,|\, \mathcal{F}_u\right] \\
&= \mathbb{E}\left[B_t^3 - B_u^3 - 3tB_u + 3uB_u \,|\, \mathcal{F}_u\right] + 0
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{E}\left[B_t^3 - B_u^3 \,|\, \mathcal{F}_u\right] - 3B_u(t-u) \\
&= \mathbb{E}\left[(B_t - B_u)^3 + 3B_t^2 B_u - 3B_u^2 B_t \,|\, \mathcal{F}_u\right] - 3B_u(t-u) \\
&= \mathbb{E}\left[(B_t - B_u)^3 \,|\, \mathcal{F}_u\right] + 3B_u \mathbb{E}\left[B_t^2 \,|\, \mathcal{F}_u\right] - 3B_u^2 \mathbb{E}[B_t \,|\, \mathcal{F}_u] - 3B_u(t-u) \\
&= \mathbb{E}\left[(B_t - B_u)^3\right] + 3B_u\left(\mathbb{E}\left[B_t^2 - t \,|\, \mathcal{F}_u\right] + t\right) - 3B_u^3 - 3B_u(t-u) \\
&= 0 + 3B_u\left(B_u^2 - u + t\right) - 3B_u^3 - 3B_u(t-u) \\
&= 0.
\end{aligned}
$$

Here we use several applications of the fact that $B_t - B_u$ is independent of $\mathcal{F}_u$, whilst B_u is $\mathcal{F}_u$ measurable. We use also that $\mathbb{E}[Z^3] = 0$ where $Z \sim N(0, \sigma^2)$, which comes from part (c) of **10.4** (or use that the normal distribution is symmetric about 0), as well as that both B_t and $B_t^2 - t$ are martingales (from Lemmas 10.4.3 and 10.4.4).

10.7 (a) We have

$$
\begin{aligned}
\sum_{k=0}^{n-1}(t_{k+1} - t_k) &= (t_n - t_{n-1}) + (t_{n-1} - t_{n-2}) + \ldots + (t_2 - t_1) + (t_1 - t_0) \\
&= t_n - t_0 \\
&= t - 0 \\
&= t.
\end{aligned}
$$

This is known as a 'telescoping sum'. The same method shows that $\sum_{k=0}^{n-1}(B_{t_{k+1}} - B_{t_k}) = B_t - B_0 = 0$.

(b) We need a bit more care for this one. We have

$$
\begin{aligned}
0 \le \sum_{k=0}^{n-1}(t_{k+1} - t_k)^2 &\le \sum_{k=0}^{n-1}(t_{k+1} - t_k)\left(\max_{j=0,\ldots,n-1}|t_{j+1} - t_j|\right) \\
&= \left(\max_{j=0,\ldots,n-1}|t_{j+1} - t_j|\right)\sum_{k=0}^{n-1}(t_{k+1} - t_k) \\
&= \left(\max_{j=0,\ldots,n-1}|t_{j+1} - t_j|\right) t.
\end{aligned}
$$

Here, the last line is deduced using part (a). Letting $n \to 0$ we have $\max_{j=0,\ldots,n-1}|t_{j+1} - t_j| \to 0$, so the right hand side of the above tends to zero as $n \to \infty$. Hence, using the sandwich rule, we have that $\sum_{k=0}^{n-1}(t_{k+1} - t_k)^2 \to 0$.

(c) Using the properties of Brownian motion, $B_{t_k+1} - B_{t_k} \sim N(0, t_{k+1} - t_k)$ so from exercise **10.4** we have

$$
\begin{aligned}
\mathbb{E}\left[\sum_{k=0}^{n-1}(B_{t_k+1} - B_{t_k})^2\right] &= \sum_{k=0}^{n-1}\mathbb{E}\left[(B_{t_k+1} - B_{t_k})^2\right] \\
&= \sum_{k=0}^{n-1}(t_{k+1} - t_k) \\
&= t.
\end{aligned}
$$

Here, the last line follows by part (a).

For the last part, the properties of Brownian motion give us that each increment $B_{t_{k+1}} - B_{t_k}$ is independent of $\mathcal{F}_{t_k}$. In particular, the increments $B_{t_{k+1}} - B_{t_k}$ are independent of each other. So, using exercise **10.4**,

$$
\begin{aligned}
\operatorname{var}\left(\sum_{k=0}^{n-1}(B_{t_k+1} - B_{t_k})^2\right) &= \sum_{k=0}^{n-1}\operatorname{var}\left((B_{t_k+1} - B_{t_k})^2\right) \\
&= \sum_{k=0}^{n-1}2(t_{k+1} - t_k)^2
\end{aligned}
$$

which tends to zero as $n \to \infty$, by the same calculation as in part (b).

10.8 (a) Let $y \geq 1$. Then

$$\begin{aligned}\mathbb{P}\left[B_t \geq y\right] &= \int_y^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2t}}\,dz \\ &\leq \int_y^\infty \frac{1}{\sqrt{2\pi t}} y e^{-\frac{z^2}{2t}}\,dz \\ &\leq \int_y^\infty \frac{1}{\sqrt{2\pi t}} z e^{-\frac{z^2}{2t}}\,dz \\ &= \frac{1}{\sqrt{2\pi t}}\left[-te^{-\frac{z^2}{2t}}\right]_{z=y}^\infty \\ &= \sqrt{\frac{t}{2\pi}} e^{-\frac{y^2}{2t}}.\end{aligned}$$

Putting $y = t^\alpha$ where $\alpha > \frac{1}{2}$ we have

$$\mathbb{P}[B_t \geq t^\alpha] \leq \sqrt{\frac{t}{2\pi}} e^{-\frac{1}{2}t^{2\alpha-1}}. \tag{C.1}$$

Since $2\alpha - 1 \geq 0$, the exponential term dominates the square root, and right hand side tends to zero as $t \to \infty$.
If $\alpha = \frac{1}{2}$ then we can use an easier method. Since $B_t \sim N(0,t)$, we have $t^{-1/2}B_t \sim N(0,1)$ and hence

$$\mathbb{P}[B_t \geq t^{1/2}] = \mathbb{P}[N(0,1) \geq 1] \in (0,1),$$

which is independent of t and hence does not tend to zero as $t \to \infty$. Note that we can't deduce this fact using the same method as for $\alpha > \frac{1}{2}$, because (C.1) only gives us an upper bound on $\mathbb{P}[B_t \geq t^\alpha]$.

(b) For this we need a different technique. For $y \geq 0$ we integrate by parts to note that

$$\begin{aligned}\int_y^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{-z^2}{2t}}\,dz &= \frac{1}{\sqrt{2\pi t}} \int_y^\infty \frac{1}{z} z e^{-\frac{-y^2}{2t}}\,dz \\ &= \frac{1}{\sqrt{2\pi t}}\left(\left[-\frac{t}{z}e^{-\frac{z^2}{2t}}\right]_{z=y}^\infty - \int_y^\infty \frac{t}{z^2}e^{-\frac{z^2}{2t}}\right)dz \\ &\leq \frac{1}{\sqrt{2\pi t}}\left(\left[-\frac{t}{z}e^{-\frac{z^2}{2t}}\right]_{z=y}^\infty\right) \\ &= \frac{1}{\sqrt{2\pi t}}\frac{t}{y}e^{-\frac{y^2}{2t}} \\ &= \frac{t}{\sqrt{2\pi}}\frac{1}{y}e^{-\frac{y^2}{2t}}\end{aligned}$$

Using the symmetry of normal random variables about 0, along with this inequality, we have

$$\begin{aligned}\mathbb{P}[|B_t| \geq a] &= \mathbb{P}[B_t \geq a] + \mathbb{P}[B_t \leq -a] \\ &= 2\mathbb{P}[B_t \geq a] \\ &\leq 2\frac{t}{\sqrt{2\pi}}\frac{1}{a}e^{-\frac{a^2}{2t}}.\end{aligned}$$

As $t \searrow 0$, the exponential term tends to 0, which dominates the $\sqrt{t}$, meaning that $\mathbb{P}[|B_t| \geq a] \to 0$ as $t \searrow 0$. That is, $B_t \to 0$ in probability as $t \searrow 0$.

Chapter 11

11.1 From (11.8) we have both $\int_0^t 1\,dB_u = B_t$ and $\int_0^s dB_u = B_s$, and using (11.6) we obtain

$$\int_u^t 1\,dB_u = \int_0^t 1\,dB_u - \int_0^v 1\,dB_u = B_t - B_v.$$

11.2 Since B_t is adapted to $\mathcal{F}_t$, we have that e^{B_t} is adapted to $\mathcal{F}_t$. Since B_t is a continuous stochastic process and $\exp(\cdot)$ is a continuous function, e^{B_t} is also continuous. From (10.2) we have that

$$\begin{aligned}\int_0^t \mathbb{E}\left[\left(e^{B_u}\right)^2\right] du &= \int_0^t \mathbb{E}\left[e^{2B_u}\right] du \\ &= \int_0^t e^{\frac{1}{2}(2^2)u}\, du \\ &= \int_0^t e^{2u}\, du < \infty.\end{aligned}$$

Therefore, $e^{B_t} \in \mathcal{H}^2$.

11.3 (a) We have

$$\mathbb{E}\left[e^{\frac{Z^2}{2}}\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{z^2}{2}} e^{\frac{-z^2}{2}}\, dz = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} 1\, dz = \infty.$$

(b) Using the scaling properties of normal distributions, $t^{-1/2}B_t$ has a $N(0,1)$ distribution for all t. Hence, if we set $F_t = t^{-1/2}B_t$ then by part (a) we have

$$\int_0^t \mathbb{E}\left[e^{\frac{1}{2}F_t^2}\right] du = \int_0^t \infty\, du$$

which is not finite. Note also that F_t is adapted to $\mathcal{F}_t$. Since e^t, t^2, $t^{-1/2}$ and B_t are all continuous, so is $e^{\frac{1}{2}F_t^2}$. Hence $e^{\frac{1}{2}F_t^2}$ is an example of a continuous, adapted stochastic process that is not in $\mathcal{H}^2$.

Note that we can't simply use the stochastic process $F_t = Z$, because we have nothing to tell us that Z is $\mathcal{F}_t$ measurable.

11.4 We have

$$\begin{aligned}\mathbb{E}[X_t] &= \mathbb{E}[2] + \mathbb{E}\left[\int_0^t t + B_u^2\, du\right] + \mathbb{E}\left[\int_0^t B_u^2\, dB_u\right] \\ &= 2 + \int_0^t t + \mathbb{E}[B_u^2]\, du + 0 \\ &= 2 + t^2 + \int_0^t u\, du \\ &= 2 + t^2 + \frac{t^2}{2} \\ &= 2 + \frac{3t^2}{2}.\end{aligned}$$

11.5 (a) $X_t = 0 + \int_0^t 0\, du + \int_0^t 0\, dB_u$ so X_t is an Ito process.

(b) $Y_t = 0 + \int_0^t 2u\, du + \int_0^t 1\, dB_u$ by (11.8), so Y_t is an Ito process.

(c) A symmetric random walk is a process in discrete time, and is therefore not an Ito process.

11.6 We have

$$\begin{aligned}\mathbb{E}[V_t] &= \mathbb{E}[e^{-kt}v] + \sigma e^{-kt}\mathbb{E}\left[\int_0^t e^{ks}\, dB_s\right] \\ &= e^{-kt}v.\end{aligned}$$

Here we use Theorem 11.2.1 to show that the expectation of the dB_u integral is zero. In order to calculate $\text{var}(V_t)$ we first calculate

$$\begin{aligned}\mathbb{E}\left[V_t^2\right] &= \mathbb{E}[e^{-2kt}v^2] + 2\sigma e^{-kt}\mathbb{E}\left[\int_0^t e^{ks}\, dB_s\right] + \sigma^2 e^{-2kt}\mathbb{E}\left[\left(\int_0^t e^{ks}\, dB_s\right)^2\right] \\ &= e^{-2kt}v^2 + 0 + \sigma^2 e^{-2kt}\int_0^t \mathbb{E}\left[(e^{ks})^2\right] du \\ &= e^{-2kt}v^2 + \sigma^2 e^{-2kt}\int_0^t e^{2ku}\, du\end{aligned}$$

$$
\begin{aligned}
&= e^{-2kt}v^2 + \sigma^2 e^{-2kt}\frac{1}{2k}\left(e^{2kt}-1\right)\\
&= e^{-2kt}v^2 + \frac{\sigma^2}{2k}\left(1-e^{-2kt}\right)\\
&=
\end{aligned}
$$

Again, we use Theorem 11.2.1 to calculate the final term on the first line. We obtain that

$$
\begin{aligned}
\mathrm{var}(V_t) &= \mathbb{E}[V_t^2] - \mathbb{E}[V_t]^2\\
&= \frac{\sigma^2}{2k}\left(1-e^{-2kt}\right).
\end{aligned}
$$

11.7 We have

$$X_t = \mu t + \int_0^t \sigma_u \, dB_u.$$

Since $M_t = \int_0^t \sigma_u \, dB_u$ is a martingale (by Theorem 11.2.1), we have that M_t is adapted and in L^1, and hence also X_t is adapted and in L^1. For $v \le t$ we have

$$
\begin{aligned}
\mathbb{E}\left[X_t \,|\, \mathcal{F}_v\right] &= \mu t + \mathbb{E}\left[M_t \,|\, \mathcal{F}_v\right]\\
&= \mu t + M_v\\
&= \mu t + \int_0^v \sigma_u \, dB_u\\
&\ge \mu v + \int_0^v \sigma_u \, dB_u\\
&= X_v.
\end{aligned}
$$

Hence, X_t is a submartingale.

11.8 (a) Taking $F_t = 0$, we have $\mathbb{E}[F_t] =$ and $\int_0^t F_s \, ds = 0$, hence $\int_0^t \mathbb{E}[F_s] \, ds = \mathbb{E}[\int_0^t F_s \, dB_s] = 0$.

(b) Taking $F_t = 1$, we have $\mathbb{E}[F_t] = 1$ and $\int_0^t F_s \, ds = t$, hence $\int_0^t \mathbb{E}[F_s] \, ds = t$ and $\mathbb{E}[\int_0^t F_s \, dB_s] = \mathbb{E}[B_t] = 0$.

11.9 (a) We have $(|X| - |Y|)^2 \ge 0$, so $2|XY| \le X^2 + Y^2$, which by monotonicity of $\mathbb{E}$ means that $2\mathbb{E}[|XY|] \le \mathbb{E}[X^2] + \mathbb{E}[Y^2]$. Using the relationship between $\mathbb{E}$ and $|\cdot|$ we have

$$2|\mathbb{E}[XY]| \le 2\mathbb{E}[|XY|] \le \mathbb{E}[X^2] + \mathbb{E}[Y^2].$$

(b) Let $X_t, Y_t \in \mathcal{H}^2$ and let $\alpha, \beta \in \mathbb{R}$ be deterministic constants. We need to show that $Z_t \alpha X_t + \beta Y_t \in \mathcal{H}^2$. Since both X_t and Y_t are continuous and adapted, Z_t is also both continuous and adapted. It remains to show that (11.5) holds for Z_t. With this in mind we note that

$$Z_t^2 = \alpha^2 X_t^2 + 2\alpha\beta X_t Y_t + \beta^2 Y_t^2$$

and hence that

$$
\begin{aligned}
|\mathbb{E}[Z_t^2]| &= \left|\alpha^2\mathbb{E}[X_t^2] + 2\alpha\beta\mathbb{E}[X_tY_t] + \beta^2\mathbb{E}[Y_t^2]\right|\\
&\le \alpha^2\mathbb{E}[X_t^2] + 2|\alpha\beta|\;|\mathbb{E}[X_tY_t]| + \beta^2\mathbb{E}[Y_t^2]\\
&\le \alpha^2\mathbb{E}[X_t^2] + |\alpha\beta|(\mathbb{E}[X_t^2] + \mathbb{E}[Y_t^2]) + \beta^2\mathbb{E}[Y_t^2]\\
&= (\alpha^2 + |\alpha\beta|)\mathbb{E}[X_t^2] + (\beta^2 + |\alpha\beta|)\mathbb{E}[Y_t^2].
\end{aligned}
$$

where we use part (a) to deduce the third line from the second. Hence,

$$
\begin{aligned}
\int_0^t \mathbb{E}[Z_u^2] \, du &\le \int_0^t (\alpha^2 + |\alpha\beta|)\mathbb{E}[X_u^2] + (\beta^2 + |\alpha\beta|)\mathbb{E}[Y_u^2] \, du\\
&= (\alpha^2 + |\alpha\beta|)\int_0^t \mathbb{E}[X_u^2] \, du + (\beta^2 + |\alpha\beta|)\int_0^t \mathbb{E}[Y_u^2] \, du < \infty
\end{aligned}
$$

as required. The final line is $< \infty$ because $X_t, Y_t \in \mathcal{H}^2$.

11.10 We have $I_F(t)=\sum\limits_{i=1}^{n} F_{t_{i-1}}[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]$. We are looking to show that

$$\mathbb{E}[I_F(t)^2]=\int_0^t \mathbb{E}[F_u^2]\,du. \tag{C.2}$$

On the right hand side we have

$$\begin{aligned}\int_0^t \mathbb{E}[F_u^2]\,du &= \sum_{i=1}^m \int_{t_{i-1}\wedge t}^{t_i\wedge t}\mathbb{E}[F_u^2]\,du\\ &= \sum_{i=1}^m (t_i\wedge t - t_{i-1}\wedge t)\mathbb{E}[F_{t_{i-1}}^2] \end{aligned} \tag{C.3}$$

because F_t is constant during each time interval $[t_{i-1}\wedge t, t_i\wedge t)$. On the left hand side of (C.2) we have

$$\begin{aligned}\mathbb{E}[I_F(t)^2] &= \mathbb{E}\left[\left(\sum_{i=1}^m F_{t_{i-1}}[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]\right)^2\right]\\ &= \mathbb{E}\left[\sum_{i=1}^m F_{t_{i-1}}^2[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]^2 + 2\sum_{i=1}^m\sum_{j=1}^{i-1} F_{t_{i-1}}F_{t_{j-1}}[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}][B_{t_j\wedge t}-B_{t_{j-1}\wedge t}]\right]\\ &= \sum_{i=1}^m \mathbb{E}\left[F_{t_{i-1}}^2[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]^2\right] + 2\sum_{i=1}^m\sum_{j=1}^{i-1}\mathbb{E}\left[F_{t_{i-1}}F_{t_{j-1}}[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}][B_{t_j\wedge t}-B_{t_{j-1}\wedge t}]\right]\end{aligned}$$

In the first sum, using the tower rule, taking out what is known, independence, and then the fact that $\mathbb{E}[B_t^2]=t$, we have

$$\begin{aligned}\mathbb{E}\left[F_{t_{i-1}}^2[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]^2\right] &= \mathbb{E}\left[\mathbb{E}\left[F_{t_{i-1}}^2[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]^2\,|\,\mathcal{F}_{t_{i-1}}\right]\right]\\ &= \mathbb{E}\left[F_{t_{i-1}}^2\mathbb{E}\left[[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]^2\,|\,\mathcal{F}_{t_{i-1}}\right]\right]\\ &= \mathbb{E}\left[F_{t_{i-1}}^2\mathbb{E}\left[[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}]^2\right]\right]\\ &= \mathbb{E}\left[F_{t_{i-1}}^2(t_i\wedge t-t_{i-1}\wedge t)\right]\end{aligned}$$

In the second sum, since $j<i$ we have $t_{j-1}\le t_i$, so using the tower rule, taking out what is known, and then the martingale property of Brownian motion, we have

$$\begin{aligned}\mathbb{E}\left[F_{t_{i-1}}F_{t_{j-1}}[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}][B_{t_j\wedge t}-B_{t_{j-1}\wedge t}]\right] &= \mathbb{E}\left[\mathbb{E}\left[F_{t_{i-1}}F_{t_{j-1}}[B_{t_i\wedge t}-B_{t_{i-1}\wedge t}][B_{t_j\wedge t}-B_{t_{j-1}\wedge t}]\,|\,\mathcal{F}_{t_{i-1}}\right]\right]\\ &= \mathbb{E}\left[F_{t_{j-1}}[B_{t_j\wedge t}-B_{t_{j-1}\wedge t}]F_{t_{i-1}}\left(\mathbb{E}\left[B_{t_i\wedge t}\,|\,\mathcal{F}_{t_{i-1}}\right]-B_{t_{i-1}\wedge t}\right)\right]\\ &= \mathbb{E}\left[F_{t_{j-1}}[B_{t_j\wedge t}-B_{t_{j-1}\wedge t}]F_{t_{i-1}}\left(B_{t_{i-1}\wedge t}-B_{t_{i-1}\wedge t}\right)\right]\\ &= 0.\end{aligned}$$

Therefore,

$$\mathbb{E}[I_F(t)^2]=\sum_{i=1}^m \mathbb{E}\left[F_{t_{i-1}}^2(t_i\wedge t-t_{i-1}\wedge t)\right]$$

which matches (C.3) and completes the proof.

Chapter 12

12.1 (a) $X_t=X_0+\int_0^t 2u\,du+\int_0^t B_u\,dB_u$.

(b) $Y_T=Y_t+\int_t^T u\,du$.

By using the fundamental theorem of calculus, we obtain that Y satisfies the differential equation $\frac{dY_t}{dt}=t$. Using equation (12.6) from Example 12.1.2, we have that

$$X_t=X_0+t^2+\frac{B_t^2}{2}-\frac{t}{2}$$

which is not differentiable because B_t is not differentiable.

12.2 We have $Z_t = f(t, X_t)$ where $f(t,x) = t^3 x$ and $dX_t = \alpha\,dt + \beta\,dB_t$. By Ito's formula,

$$\begin{aligned} dZ_t &= \left\{3t^2 X_t + (\alpha)(t^3) + \frac{1}{2}(\beta)(0)\right\} dt + \beta t^3\,dB_t \\ &= \left(3t^2 X_t + \alpha t^3\right) dt + \beta t^3\,dB_t. \end{aligned}$$

12.3 By Ito's formula, using that $dB_t = 0\,dt + 1\,dB_t$, we have

(a) $dZ_t = \left\{B_t^2 + (0)(2tB_t) + \frac{1}{2}(1^2)(2t)\right\} dt + (1)(2tB_t)\,dB_t = \left(B_t^2 + t\right) dt + 2tB_t\,dB_t.$

(b) $dZ_t = \left\{\alpha e^{\alpha t} + (0)(0) + \frac{1}{2}(1^2)(0)\right\} dt + (1)(0)\,dB_t = \alpha e^{\alpha t}\,dt.$

(c) We have

$$\begin{aligned} dZ_t &= \left\{(0) + (t^2)\left(\frac{-1}{X_t^2}\right) + \frac{1}{2}(B_t)^2\left(\frac{2}{X_t^3}\right)\right\} dt + (B_t)\left(\frac{-1}{X_t^2}\right) B_t \\ &= \left(B_t^2 Z_t^3 - t^2 Z_t^2\right) dt - B_t Z_t^2\,dB_t. \end{aligned}$$

(d) We have

$$\begin{aligned} dZ_t &= \left\{(0) + (\cos X_t)\,(\cos X_t) + \frac{1}{2}(\cos X_t)^2\,(-\sin X_t)\right\} dt + (\cos X_t)(\cos X_t)\,dB_t \\ &= \left(1 - \tfrac{Z_t}{2}\right)\left(1 - Z_t^2\right) dt + \left(1 - Z_t^2\right) dB_t \end{aligned}$$

12.4 We have

$$\begin{aligned} dF_t &= \left(0 + (0)(nB_t^{n-1}) + \frac{1}{2}(1^2)(n(n-1)B_t^{n-2}\right) dt + (1)nB_t^{n-1}\,dB_t \\ &= \frac{n(n-1)B_t^{n-2}}{2}\,dt + nB_t^{n-1}\,dB_t. \end{aligned}$$

Written out in integral form this gives

$$B_n^t = \int_0^t \frac{n(n-1)B_u^{n-2}}{2}\,du + \int_0^t nB_u^{n-1}\,dB_u.$$

Taking expectations, swapping $\int du$ with $\mathbb{E}$, and recalling from Theorem 11.2.1 that integrals with respect to dB_t have zero mean, we obtain

$$\begin{aligned} \mathbb{E}[B_t^n] &= \int_0^t \frac{n(n-1)\mathbb{E}[B_u^{n-2}]}{2}\,du + 0 \\ &= \frac{n(n-1)}{2}\int_0^t \mathbb{E}[B_u^{n-2}]\,du. \end{aligned}$$

12.5 (a) By Ito's formula, with $f(t,x) = e^{t/2}\cos x$, we have

$$\begin{aligned} dX_t &= \left(\frac{1}{2}e^{t/2}\cos(B_t) + (0)\left(-e^{t/2}\sin(B_t)\right) + \frac{1}{2}(1^2)(-e^{t/2}\cos(B_t))\right) dt + (1)\left(-e^{t/2}\sin(B_t)\right) dB_t \\ &= -e^{t/2}\sin(B_t)\,dB_t. \end{aligned}$$

Hence $X_t = X_0 - \int_0^t e^{u/2}\sin(B_u)\,dB_u$, which is martingale by Theorem 11.2.1.

(b) By Ito's formula, with $f(t,x) = (x+t)e^{-x-t/2}$ we have

$$\begin{aligned} dY_t &= \Big\{e^{-B_t - t/2} - \frac{1}{2}(B_t+t)e^{-B_t-t/2} + (0)\left(e^{-B_t-t/2} - (B_t+t)e^{-B_t-t/2}\right) \\ &\qquad + \frac{1}{2}(1^2)\left(-e^{-B_t-t/2} - e^{-B_t-t/2} + (B_t+t)e^{-B_t-t/2}\right)\Big\} dt \\ &\qquad + (1^2)\left(e^{-B_t-t/2} - (B_t+t)e^{-B_t-t/2}\right) dB_t \\ &= (1 - t - B_t)e^{-B_t-t/2}\,dB_t. \end{aligned}$$

Hence $Y_t = Y_0 + \int_0^t (1 - u - B_u)e^{-B_u-u/2}\,dB_u$, which is martingale by Theorem 11.2.1.

12.6 We apply Ito's formula with $f(t,x)=tx$ to $Z_t=f(t,B_t)=tB_t$ and obtain

$$dZ_t=\left(B_t+(0)(t)+\frac{1}{2}(1^2)(0)\right)\,dt+(1)(t)\,dB_t$$

so we obtain

$$tB_t=0B_0+\int_0^t B_u\,du+\int_0^t u\,dB_u,$$

as required.

12.7 (a) We have $X_t=X_0+\int_0^t 2+2s\,ds+\int_0^t B_s\,dB_s$. Taking expectations, and recalling that Ito integrals have zero mean, we obtain that

$$\mathbb{E}[X_t]=X_0+\int_0^t 2+2s\,ds+0=1+\left[2s+s^2\right]_{s=0}^t=1+2t+t^2=(1+t)^2.$$

(b) From Ito's formula,

$$\begin{aligned}dY_t&=\left(0+(2+2t)(2X_t)+\tfrac{1}{2}(B_t)^2(2)\right)\,dt+B_t(2X_t)\,dB_t\\&=\left(4(1+t)X_t+(B_t)^2\right)\,dt+2X_tB_t\,dB_t.\end{aligned}$$

Writing in integral form, taking expectations, and using that Ito integrals have zero mean, we obtain

$$\begin{aligned}\mathbb{E}[Y_t]&=Y_0+\mathbb{E}\left[\int_0^t 4(1+s)X_t+(B_s)^2\,ds\right]+0\\&=1+\int_0^t 4(1+s)\mathbb{E}[X_s]+\mathbb{E}[B_s^2]\,ds\\&=1+\int_0^t 4(1+s)^3+s\,ds\\&=1+\left[(1+s)^4+\tfrac{s^2}{2}\right]_{s=0}^t\\&=(1+t)^4+\tfrac{t^2}{2}\end{aligned}$$

Hence, using that $\mathbb{E}[X_t^2]=\mathbb{E}[Y_t]$,

$$\begin{aligned}\mathrm{var}(X)&=\mathbb{E}[X_t^2]-\mathbb{E}[X_t]^2\\&=(1+t)^4+\tfrac{t^2}{2}-(1+t)^4\\&=\tfrac{t^2}{2}.\end{aligned}$$

(c) If we change the dB_t coefficient then we won't change the mean, because we can see from (a) that $\mathbb{E}[X_t]$ depends only on the dB_t coefficient. However, as we can see from part (b), the variance depends on both the dt and dB_t terms, so will typically change if we alter the dB_t coefficient.

12.8 In integral form, we have

$$X_t=X_0+\int_0^t \alpha X_u\,du+\int_0^t \sigma_u\,dB_u.$$

Taking expectations, swapping $\int du$ with $\mathbb{E}$, and recalling from Theorem 11.2.1 that Ito integrals are zero mean martingales, we obtain

$$\mathbb{E}[X_t]=\mathbb{E}[X_0]+\int_0^t \alpha\mathbb{E}[X_u]\,du.$$

Applying the fundamental theorem of calculus, if we set $x_t=\mathbb{E}[X_t]$, we obtain

$$\frac{dx_t}{dt}=\alpha x_t$$

which has solution $x_t=Ce^{\alpha t}$. Putting in $t=0$ shows that $C=\mathbb{E}[X_0]=1$, hence

$$\mathbb{E}[X_t]=e^{\alpha t}.$$

12.9 We have $X_t = X_0 + \int_0^t X_s\,dB_s$. Since Ito integrals are zero mean martingales, this means that $\mathbb{E}[X_t] = \mathbb{E}[X_0] = 1$. Writing $Y_t = X_t^2$ and using Ito's formula,

$$\begin{aligned} dY_t &= \left(0 + (0)(2X_t) + \tfrac{1}{2}(X_t)^2(2)\right)\,dt + (X_t)(2X_t)\,dB_t \\ &= (X_t^2)\,dt + 2X_t^2\,dB_t. \\ &= Y_t\,dt + 2Y_t\,dB_t \end{aligned}$$

Writing in integral form and taking expectations, we obtain

$$\mathbb{E}[Y_t] = 1 + \int_0^t \mathbb{E}[Y_s]\,ds + 0.$$

Hence, by the fundamental theorem of calculus, $f(t) = \mathbb{E}[Y_t]$ satisfies the differential equation $f'(t) = f(t)$. The solution of this differential equation is $f(t) = Ae^t$. Since $\mathbb{E}[Y_0] = 1$ we have $A = 1$ and thus $\mathbb{E}[Y_t] = e^t$. Hence,

$$\mathrm{var}(X_t) = \mathbb{E}[X_t^2] - \mathbb{E}[X_t]^2 = \mathbb{E}[Y_t] - 1 = e^t - 1.$$

12.10 (a) Applying Ito's formula to $Z_t = B_t^3$, with $f(t,x) = x^3$, we have

$$\begin{aligned} dZ_t &= \left(0 + (0)(3B_t^2) + \frac{1}{2}(1^2)(6B_t)\right)\,dt + (1)(3B_t^2)\,dB_t \\ &= 3B_t\,dt + 3B_t^2\,dt \end{aligned}$$

and substituting in for Z we obtain

$$dZ_t = 3Z_t^{1/3}\,dt + 3Z_t^{2/3}\,dB_t$$

as required.

(b) Another solution is the (constant, deterministic) solution $X_t = 0$.

12.11 Equation (12.9) says that

$$X_t = X_0 \exp\left((\alpha - \tfrac{1}{2}\sigma^2)t + \sigma B_t\right).$$

Using Ito's formula, with $f(t,x) = X_0\exp\left((\alpha - \frac{1}{2}\sigma^2)t + \sigma x\right)$ we obtain

$$\begin{aligned} dX_t &= \left((\alpha - \tfrac{1}{2}\sigma^2)X_t + (0)(\sigma X_t) + \frac{1}{2}(1^2)(\sigma^2 X_t)\right)\,dt + (1)(\sigma X_t)dB_t \\ &= \alpha X_t\,dt + \sigma X_t\,dB_t \end{aligned}$$

and thus X_t solves (12.8).

12.12 Equation (12.12) says that

$$X_t = X_0 \exp\left(\int_0^t \sigma_u\,dB_u - \frac{1}{2}\int_0^t \sigma_u^2\,du\right)$$

We need to arrange this into a form where we can apply Ito's formula. We write

$$X_t = X_0 \exp\left(Y_t - \frac{1}{2}\int_0^t \sigma_u^2\,du\right)$$

where $dY_t = \sigma_t\,dB_t$ with $Y_0 = 0$. We now have $X_t = f(t, Y_t)$ where $f(t,y) = X_0\exp\left(y - \frac{1}{2}\int_0^t \sigma_u\,du\right)$, so from Ito's formula (and the fundamental theorem of calculus) we obtain

$$\begin{aligned} dX_t &= \left(-\frac{1}{2}\sigma_t^2 X_t + (0)(X_t) + \frac{1}{2}(\sigma_t^2)(X_t)\right)\,dt + (\sigma_t)(X_t)\,dB_t \\ &= \sigma_t X_t\,dB_t. \end{aligned}$$

Hence, X_t solves (12.11).

12.13 (a) This is essentially Example 3.3.9 but in continuous time. By definition of conditional expectation (i.e. Theorem 3.1.1) we have that $M_t \in L^1$ and that $M_t \in \mathcal{F}_t$. It remains only to use the tower property to note that for $0 \le u \le t$ we have

$$\mathbb{E}[M_t|\mathcal{F}_u] = \mathbb{E}[\mathbb{E}[Y\,|\,\mathcal{F}_t]\mathcal{F}_u] = \mathbb{E}[Y\,|\,\mathcal{F}_u] = M_u.$$

(b) (i) Note that $M_0 = \mathbb{E}[B_T^2 \,|\, \mathcal{F}_0] = \mathbb{E}[B_T^2] = T$. We showed in Lemma 10.4.4 that $B_t^2 - t$ was a martingale, hence

$$\begin{aligned}\mathbb{E}[B_T^2 \,|\, \mathcal{F}_t] &= \mathbb{E}[B_T^2 - T + T \,|\, \mathcal{F}_t] \\ &= B_t^2 - t + T\end{aligned}$$

Using (12.6), this gives us that

$$\mathbb{E}[B_T^2 \,|\, \mathcal{F}_t] = 2\int_0^t B_u \, dB_u + t - t + T$$

so we obtain

$$M_t = T + \int_0^t 2B_u \, dB_u$$

and we can take $h_t = 2B_t$.

(ii) Note that $M_0 = \mathbb{E}[B_T^3 \,|\, \mathcal{F}_0] = \mathbb{E}[B_T^3] = 0$. We showed in **10.6** that $B_t^3 - 3tB - t$ was a martingale. Hence,

$$\begin{aligned}\mathbb{E}[B_T^3 \,|\, \mathcal{F}_t] &= \mathbb{E}[B_T^3 - 3TB_T + 3TB_T \,|\, \mathcal{F}_t] \\ &= B_t^3 - 3tB_t + 3TB_t.\end{aligned}$$

Using Ito's formula on $Z_t = B_t^3$, we obtain $dZ_t = \left\{0 + (0)(3B_t^2) + \frac{1}{2}(1^2)(6B_t)\right\} dt + (1)(3B_t^2)\, dB_t$ so as

$$B_t^3 = 0 + \int_0^t 3B_u \, du + \int_0^t 3B_u^2 \, dB_u.$$

Also, from **12.6** we have

$$tB_t = \int_0^t B_u \, du + \int_0^t u \, dB_u,$$

so as

$$\begin{aligned}\mathbb{E}[B_T^3 \,|\, \mathcal{F}_t] &= 3\int_0^t B_u \, du + 3\int_0^t B_u^2 \, dB_u - 3\left(\int_0^t B_u \, du + \int_0^t u \, dB_u\right) + 3TB_t \\ &= 3\int_0^t B_u^2 \, dB_u - 3\int_0^t u \, dB_u + 3T\int_0^t 1 \, dB_u \\ &= \int_0^t 3B_u^2 - 3u + 3T \, dB_u.\end{aligned}$$

We can take $h_t = 3B_t^2 - 3t + 3T$.

(iii) Note that $M_0 = \mathbb{E}[e^{\sigma B_T} \,|\, \mathcal{F}_0] = \mathbb{E}[e^{\sigma B_T}] = e^{\frac{1}{2}\sigma^2 T}$ by (10.2) and the scaling properties of normal random variables. We showed in **10.6** that $e^{\sigma B_t - \frac{1}{2}\sigma^2 t}$ was a martingale. Hence,

$$\begin{aligned}\mathbb{E}[e^{\sigma B_T} \,|\, \mathcal{F}_t] &= \mathbb{E}[e^{\sigma B_T - \frac{1}{2}\sigma^2 T} e^{\frac{1}{2}\sigma^2 T} \,|\, \mathcal{F}_t] \\ &= e^{\sigma B_t - \frac{1}{2}\sigma^2 t} e^{\frac{1}{2}\sigma^2 T} \\ &= e^{\sigma B_t - \frac{1}{2}\sigma^2 (T-t)}.\end{aligned}$$

Applying Ito's formula to $Z_t = e^{\sigma B_t - \frac{1}{2}\sigma^2(T-t)}$ gives that

$$\begin{aligned}dZ_t &= \left(-\frac{1}{2}\sigma^2 Z_t + (0)(\sigma Z_t) + \frac{1}{2}(1^2)(\sigma^2 Z_t)\right) dt + (1)(\sigma Z_t)\, dB_t \\ &= \sigma Z_t dB_t\end{aligned}$$

so we obtain that

$$Z_t = Z_0 + \int_0^t \sigma Z_u \, dB_u.$$

Substituting in for Z_t we obtain

$$\mathbb{E}[e^{\sigma B_T} \,|\, \mathcal{F}_t] = e^{\frac{1}{2}\sigma^2 T} + \int_0^t \sigma Z_u \, dB_u$$

so we can take $h_t = \sigma Z_t = \sigma e^{\sigma B_t - \frac{1}{2}\sigma^2(T-t)}$.

Chapter 13

13.1 We have $\alpha(t,x)=-2t$, $\beta=0$ and $\Phi(x)=e^x$. By Lemma 13.1.2 the solution is given by

$$F(t,x)=\mathbb{E}_{t,x}[e^{X_T}]$$

where $dX_u=-2u\,du+0\,dB_u=du$. This gives

$$\begin{aligned}X_T&=X_t-\int_t^T 2u\,du\\&=X_t-T^2+t^2.\end{aligned}$$

Hence,

$$\begin{aligned}F(t,x)&=\mathbb{E}_{t,x}\left[e^{X_t-t^2}\right]\\&=\mathbb{E}\left[e^{x-T^2+t^2}\right]\\&=e^xe^{-T^2+t^2}.\end{aligned}$$

Here we use that $X_t=x$ under $\mathbb{E}_{t,x}$.

13.2 We have $\alpha=\beta=1$ and $\Phi(x)=x^2$. By Lemma 13.1.2 the solution is given by

$$F(t,x)=\mathbb{E}_{t,x}\left[X_T^2\right]$$

where $dX_u=du+dB_u$. This gives

$$\begin{aligned}X_T&=X_t+\int_t^T du+\int_t^T dB_u\\&=X_t+(T-t)+(B_T-B_t).\end{aligned}$$

Hence,

$$\begin{aligned}F(t,x)&=E_{t,x}[(X_t+(T-t)+(B_T-B_t))^2]\\&=\mathbb{E}\left[x^2+(T-t)^2+(B_T-B_t)^2+2x(T-t)+2x(B_T-B_t)+2(T-t)(B_T-B_t)\right]\\&=x^2+(T-t)^2+(T-t)+2x(T-t).\end{aligned}$$

Here we use that $X_t=x$ under $\mathbb{E}_{t,x}$ and that $B_T-B_t\sim B_{T-t}\sim N(0,T-t)$.

13.3 (a) We have $Z_t=F(t,X_t)+\gamma(t)$, where $dX_t=\alpha(t,x)\,dt+\beta(t,x)\,dB_t$, so

$$\begin{aligned}dZ_t&=\left(\frac{\partial F}{\partial t}+\frac{\partial\gamma}{\partial t}(t)+\alpha(t,x)\frac{\partial F}{\partial x}+\frac{1}{2}\beta(t,x)^2\frac{\partial^2F}{\partial x^2}\right)dt+\beta(t,x)\frac{\partial F}{\partial x}\,dB_t\\&=\beta(t,x)\frac{\partial F}{\partial x}\,dB_t\end{aligned}$$

where as usual we have suppressed the (t,X_t) arguments of F and its partial derivatives. Note that the term in front of the dt is zero because F satisfies (13.8).

(b) We use the same strategy as in the proof of Lemma 13.1.2. Writing out dZ_t in integral form over time interval $[t,T]$ we obtain

$$Z_T=Z_t+\int_t^T\beta(u,x)\frac{\partial F}{\partial x}\,dB_u.$$

Taking expectations $\mathbb{E}_{t,x}$ gives us

$$\mathbb{E}_{x,t}\left[F(T,X_T)+\gamma(T)\right]=\mathbb{E}_{t,x}[F(t,X_t)+\gamma(t)]$$

and noting that $X_t=x$ under $\mathbb{E}_{t,x}$, we have

$$\mathbb{E}_{x,t}\left[F(T,X_T)\right]+\gamma(T)=F(t,x)+\gamma(t).$$

Using (13.9) we have $F(T,X_T)=\Phi(X_T)$ and we obtain

$$\mathbb{E}_{x,T}[\Phi(X_T)]+\gamma(T)-\gamma(t)=F(t,x)$$

as required.

13.4 Consider (for example) the stochastic process

$$M_t = \begin{cases} B_t & \text{for } t \le 1, \\ B_t - B_{t-2} & \text{for } t > 1. \end{cases}$$

We now consider M_t at time 3. The intuition is that $\mathbb{E}[M_3 \,|\, \mathcal{F}_2]$ can see the value of B_1, but that $\sigma(M_2) = \sigma(B_2)$ and $\mathbb{E}[M_3 \,|\, \sigma(M_2)]$ cannot see the value of B_1.
Formally: we have

$$\begin{aligned} \mathbb{E}[M_3 \,|\, \mathcal{F}_2] &= \mathbb{E}[B_3 - B_1 \,|\, \mathcal{F}_2] \\ &= B_2 - B_1 \end{aligned} \tag{C.4}$$

Here, we use that B_t is a martingale. However,

$$\begin{aligned} \mathbb{E}_{2,M_2}[M_3] &= \mathbb{E}[B_3 - B_1 \,|\, \sigma(M_2)] \\ &= \mathbb{E}[B_3 - B_1 \,|\, \sigma(B_2 - B_0)] \\ &= \mathbb{E}[B_3 - B_1 \,|\, \sigma(B_2)] \\ &= \mathbb{E}[B_3 - B_2 \,|\, \sigma(B_2)] + \mathbb{E}[B_2 \,|\, \sigma(B_2)] - \mathbb{E}[B_1 \,|\, \sigma(B_2)] \\ &= \mathbb{E}[B_3 - B_2] + B_2 - \mathbb{E}[B_1 \,|\, \sigma(B_2)] \\ &= 0 + B_2 - \mathbb{E}[B_1 \,|\, \sigma(B_2)] \\ &= B_2 - \mathbb{E}[B_1 \,|\, \sigma(B_2)] \end{aligned} \tag{C.5}$$

If we can show that (C.4) and (C.5) are not equal, then we have that M_t is not Markov. Their difference is $D = \text{(C.4)} - \text{(C.5)} = \mathbb{E}[B_1 \,|\, \sigma(B_2)] - B_1$.
We can write $B_2 = (B_2 - B_1) + (B_1 - B_0)$. By the properties of Brownian motion, $B_2 - B_1$ and $B_1 - B_0$ are independent and identically distributed. Hence, by symmetry, $\mathbb{E}[B_2 - B_1 \,|\, \sigma(B_2)] = \mathbb{E}[B_1 - B_0 \,|\, \sigma(B_2)]$ and since

$$B_2 = \mathbb{E}[B_2 \,|\, \sigma(B_2)] = \mathbb{E}[B_2 - B_1 \,|\, \sigma(B_2)] + \mathbb{E}[B_1 - B_0 \,|\, \sigma(B_2)]$$

we have that $\mathbb{E}[B_2 - B_1 \,|\, \sigma(B_2)] = \frac{B_2}{2}$. Hence

$$D = \frac{B_2}{2} - B_1$$

which is non-zero.

Chapter 14

14.1 We have

$$\frac{\partial f}{\partial t} = 0, \quad \frac{\partial f}{\partial s} = c, \quad \frac{\partial^2 f}{\partial x^2} = 0$$

which, put into (14.10), gives $(0) + rs(c) + \frac{1}{2}s^2\sigma^2(0) - r(cs) = 0$. Similarly,

$$\frac{\partial g}{\partial t} = rce^{rt}, \quad \frac{\partial g}{\partial s} = 0, \quad \frac{\partial^2 g}{\partial x^2} = 0$$

which, put into (14.10), gives $(rce^{rt}) + rs(0) + \frac{1}{2}s^2\sigma^2(0) - r(ce^{rt}) = 0$.

14.2 We have

$$\begin{aligned} e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_T) \,|\, \mathcal{F}_t\right] &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[K\right] \\ &= Ke^{-r(T-t)} \end{aligned}$$

because K is deterministic. By Theorem 14.3.1 this is the price of the contingent claim $\Phi(S_T)$ at time t.
We can hedge the contingent claim K simply by holding Ke^{-rT} cash at time 0, and then waiting until time T. While we wait, the cash increases in value according to (14.2) i.e. at continuous rate r. So, this answer is not surprising.

14.3 (a) By Theorem 14.3.1 the price of the contingent claim $\Phi(S_T)=\log(S_T)$ at time t is

$$e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_T)\,|\,\mathcal{F}_t\right]=e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\log(S_T)\,|\,\mathcal{F}_t\right].$$

Recall that, under $\mathbb{Q}$, S_t is a geometric Brownian motion, with $S_0=0$, drift r and volatility σ. So from (14.20) we have $S_T=S_te^{(r-\frac12\sigma^2)t+\sigma B_t}$. Hence,

$$\begin{aligned}e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\log(S_T)\,|\,\mathcal{F}_t\right]&=e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\log(S_t)+(r-\tfrac12\sigma^2)(T-t)+\sigma(B_T-B_t)\,|\,\mathcal{F}_t\right]\\&=e^{-r(T-t)}\left(\log(S_t)+(r-\tfrac12\sigma^2)(T-t)+\sigma\left(\mathbb{E}^{\mathbb{Q}}[B_T\,|\,\mathcal{F}_t]-B_t\right)\right)\\&=e^{-r(T-t)}\left(\log(S_t)+(r-\tfrac12\sigma^2)(T-t)\right).\end{aligned}$$

Here, we use that S_t,B_t are $\mathcal{F}_t$ measurable, and that (B_t) is a martingale.
Putting in $t=0$ we obtain

$$e^{-rT}\left(\log(S_0)+(r-\tfrac12\sigma^2)T\right),$$

(b) The strategy could be summarised as 'write down $\Phi(S_T)$, replace T with $t=0$ and hope'. The problem is that if we buy $\log s$ units of stock at time 0, then from (14.20) (with $t=0$) its value at time T will be $(\log s)\exp((r-\frac12\sigma^2)T+\sigma B_T)$, which is not equal to $\log S_T=\log s+(r-\frac12\sigma^2)T+\sigma B_T$.
(In more formal terminology, the problem is that our excitable mathematician has assumed, incorrectly, that the log *function and the 'find the price' function commute with each other.)*

14.4 (a) We have (in the risk neutral-world $\mathbb{Q}$) that $dS_t=rS_t\,dt+\sigma S_t\,dB_t$. Hence, by Ito's formula,

$$\begin{aligned}dY_t&=\left((0)+rS_t(\beta S_t^{\beta-1})+\frac12\sigma^2S_t^2(\beta(\beta-1)S_t^{\beta-2})\right)dt+\sigma S_t(\beta S_t^{\beta-1})dB_t\\&=\left(r\beta+\tfrac12\sigma^2\beta(\beta-1)\right)Y_t\,dt+(\sigma\beta)Y_t\,dB_t.\end{aligned}$$

So Y_t is a geometric Brownian motion with drift $r\beta+\frac12\sigma^2\beta(\beta-1)$ and volatility $\sigma\beta$.

(b) Applying (14.20) and replacing the drift and volatility with those from part (a), we have that

$$\begin{aligned}Y_T&=Y_t\exp\left(\left(r\beta+\tfrac12\sigma^2\beta(\beta-1)-\tfrac12\sigma^2\beta^2\right)(T-t)+\sigma\beta(B_T-B_t)\right)\\&=Y_t\exp\left(\left(r\beta-\tfrac12\sigma^2\beta\right)(T-t)+\sigma\beta(B_T-B_t)\right).\end{aligned}$$

By Theorem 14.3.1, the arbitrage free price of the contingent claim $Y_t=\Phi(S_T)$ at time t is

$$\begin{aligned}e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[Y_T\,|\,\mathcal{F}_t\right]&=e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[S_t^\beta\exp\left(\left(r\beta-\tfrac12\sigma^2\beta\right)(T-t)+\sigma\beta(B_T-B_t)\right)\,|\,\mathcal{F}_t\right]\\&=e^{-r(T-t)}S_t^\beta e^{(r\beta-\frac12\sigma^2\beta)(T-t)}\mathbb{E}^{\mathbb{Q}}\left[e^{\sigma\beta(B_T-B_t)}\,|\,\mathcal{F}_t\right]\\&=e^{-r(T-t)}S_t^\beta e^{(r\beta-\frac12\sigma^2\beta)(T-t)}\mathbb{E}^{\mathbb{Q}}\left[e^{\sigma\beta(B_T-B_t)}\right]\\&=e^{-r(T-t)}S_t^\beta e^{(r\beta-\frac12\sigma^2\beta)(T-t)}e^{\frac12\sigma^2\beta^2(T-t)}\\&=S_t^\beta e^{-r(T-t)(1-\beta)-\frac12\sigma^2\beta(T-t)(1-\beta)}.\end{aligned}$$

Here, we use that S_t is $\mathcal{F}_t$ measurable. We then use (10.2) along with the properties of Brownian motion to tell us that $\sigma\beta(B_T-B_t)$ is independent of $\mathcal{F}_t$ with distribution $N(0,\sigma^2\beta^2(T-t))$.

14.5 From Theorem 14.3.1 the price at time t of the binary option is

$$e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S_T)\,|\,\mathcal{F}_t\right]=Ke^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{S_T\in[\alpha,\beta]\}}\,|\,\mathcal{F}_t\right]$$

We have

$$\begin{aligned}S_T&=S_te^{(r-\frac12\sigma^2)(T-t)+\sigma(B_T-B_t)}\\&=S_te^{(r-\frac12\sigma^2)(T-t)+\sigma\sqrt{T-t}Z}\end{aligned}$$

where $Z\sim N(0,1)$ is independent of $\mathcal{F}_t$. Here we use that $B_T-B_t\sim N(0,T-t)\sim\sqrt{T-t}N(0,1)$ is independent of $\mathcal{F}_t$. Therefore,

$$\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{S_T\in[\alpha,\beta]\}}\,|\,\mathcal{F}_t\right]=\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\left\{S_te^{(r-\frac12\sigma^2)(T-t)+\sigma\sqrt{T-t}Z}\in[\alpha,\beta]\right\}}\,\middle|\,\mathcal{F}_t\right]$$

$$
\begin{aligned}
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}\left\{\frac{\log\left(\frac{\alpha}{S_t}\right)-\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \le Z \le \frac{\log\left(\frac{\beta}{S_t}\right)-\left(r+\frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{T-t}}\right\}\,\middle|\,\mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}\left\{\frac{\log\left(\frac{\alpha}{S_t}\right)-\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \le Z \le \frac{\log\left(\frac{\beta}{S_t}\right)-\left(r+\frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{T-t}}\right\}\right] \\
&= \mathcal{N}(e_1)-\mathcal{N}(e_2)
\end{aligned}
$$

where

$$
\begin{aligned}
e_1 &= \frac{\log\left(\frac{\alpha}{S_t}\right)-\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \\
e_2 &= \frac{\log\left(\frac{\beta}{S_t}\right)-\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.
\end{aligned}
$$

Hence the price at time t is given by $Ke^{r(T-t)}\left[\mathcal{N}(e_1)-\mathcal{N}(e_2)\right]$.

14.6 If we take $F_t = e^{\mu t}$ then (in world $\mathbb{P}$) we have

$$\frac{S_t}{F_t} = \frac{\exp\left(\left(\mu-\frac{1}{2}\sigma^2\right)t+\sigma B_t\right)}{\exp(\mu t)} = \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right),$$

which we showed was a martingale in **10.6**.

14.7 We have the risk neutral pricing formula $\Pi_t(\Phi) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,S_t}[\Phi(S_T)]$ for any contingent claim Φ. Hence, using linearity of $\mathbb{E}^{\mathbb{Q}}$,

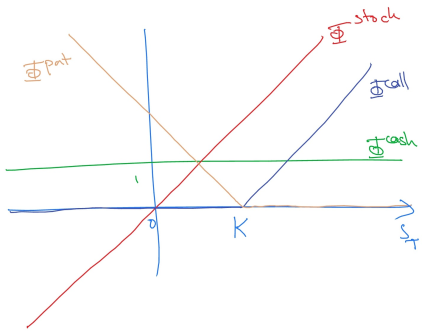
$$
\begin{aligned}
\Pi_t(\alpha\Phi_1+\beta\Phi_2) &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,S_t}\left[\alpha\Phi_1(S_T)+\beta\Phi_2(S_T)\right] \\
&= \alpha e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,S_t}\left[\Phi_1(S_T)\right]+\beta e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_{t,S_t}\left[\Phi_2(S_T)\right] \\
&= \alpha\Pi_t(\Phi_1)+\beta\Pi_t(\Phi_2)
\end{aligned}
$$

as required.

14.8 In the 'new' model we could indeed buy a unit of $Y_t = S_t^\beta$ and hold onto it for as long as we liked. But in the 'old' model we can only buy (linear multiples of) the stock S_t; we can't buy a commodity whose price at time t is S_t^β. This means that the hedging strategy suggested within the 'new' model doesn't work within the 'old' model. Consequently, there is no reason to expect that prices in the two models will be equal. In general they will not be.

Chapter 15

15.1 (a) The functions $\Phi^{cash}(S_T)=1$, $\Phi^{stock}(S_T)=S_T$, $\Phi^{call}(S_T)=\max(S_T-K,0)$ and $\Phi^{put}(S_T)=\max(K-S_T,0)$ look like:



(b) In terms of functions, the put-call parity relation states that

$$\max(K-S_T,0) = \max(S_T-K)+K-S_T.$$

To check that this holds we consider two cases.

- If $K \le S_T$ then put-call parity states that $0 = S_T-K+K-S_T$, which is true.

- If $K \geq S_T$ then put-call parity states that $K - S_T = 0 + K - S_T$, which is true.

15.2 The put-call parity relation (15.1) says that

$$\Phi^{put}(S_T) = \Phi^{call}(S_T) + K\Phi^{cash}(S_T) - \Phi^{stock}(S_T).$$

Hence,

$$\Pi_t^{put} = \Pi_t^{call} + K\Pi_t^{cash} - \Pi_t^{stock}.$$

From here, using that $\Pi_t^{cash} = e^{-r(T-t)}$ (corresponding to $\Phi(S_T) = 1$) and $\Pi_t^{stock} = S_t$ (corresponding to $\Phi(S_T) = S_T$), as well as the Black-Scholes formula (14.23), we have

$$\begin{aligned}
\Pi_t^{put} &= S_t\mathcal{N}[d_1] - Ke^{-r(T-t)}\mathcal{N}[d_2] + Ke^{-r(T-t)} - S_t \\
&= S_t(\mathcal{N}[d_1] - 1) - Ke^{-r(T-t)}(\mathcal{N}[d_2] - 1) \\
&= -S_t\mathcal{N}[-d_1] + Ke^{-r(T-t)}\mathcal{N}[-d_2].
\end{aligned}$$

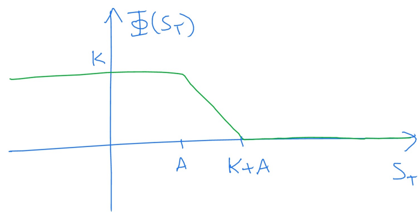
In the last line we use that $\mathcal{N}[x] + \mathcal{N}[-x] = 1$, which follows from the fact that the $N(0,1)$ distribution is symmetric about 0 (i.e. $\mathbb{P}[X \leq x] + \mathbb{P}[X \leq -x] = \mathbb{P}[X \leq x] + \mathbb{P}[-X \geq x] = \mathbb{P}[X \leq x] + \mathbb{P}[X \geq x] = 1$). The formula stated in the question follows from setting $t = 0$.

15.3 Write $\Phi^{call,K}(S_T) = \max(S_T - K, 0)$ for the contingent claim of a European call option, and $\Phi^{put,K}(S_T)$ for the contingent claim of a European put options, both with strike price K and exercise date T. Then

$$\Phi(S_T) = \Phi^{call,1}(S_T) + \Phi^{put,-1}(S_T)$$

so we can hedge $\Phi(S_T)$ by holding a single call option with strike price 1 and a single put option with strike price -1.

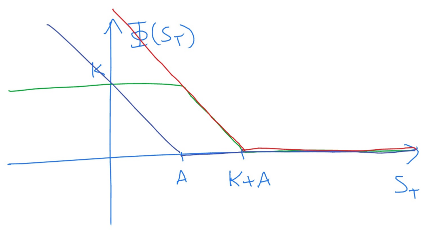
15.4 (a) A sketch of $\Phi(S_T)$ for general A looks like



Let $\Phi^{put,K}(S_T) = \max(K - S_T, 0)$ denote the contingent claim corresponding to a European put option with strike price K and exercise date T. Then, we claim that

$$\Phi(S_T) = \Phi^{put,K+A}(S_T) - \Phi^{put,A}(S_T). \tag{C.6}$$

The right-hand side of this equation corresponds to a portfolio of one put option with strike price $K + A$ and minus one put option with strike price A. We can see that the relation (C.6) holds by using a diagram



in which the purple line $\Phi^{put,A}(S_T)$ is subtracted from the red line $\Phi^{put,K+A}(S_T)$ to obtain the green line $\Phi(S_T)$. Alternatively, we can check it by considering three cases:

- If $S_T \leq A$ then we have $K = (K + A - S_T) - (A - S_T)$ which is true.
- If $A \leq S_T \leq K + A$ then we have $K + A - S_T = (K + A - S_T) - (0)$ which is true.
- If $S_T \geq K + A$ then we have $0 = (0) - (0)$ which is true.

Therefore, the portfolio of one put option with strike price $K + A$ and minus one put option with strike price A is a replicating portfolio for $\Phi(S_T)$.

(b) We use put-call parity to replicate our portfolio of put options with a portfolio of cash, stock and call options. This tells us that $\Phi^{put,K}(S_T) = \Phi^{call,K}(S_T) + K\Phi^{cash}(S_T) + \Phi^{stock}(S_T)$. Therefore, replacing each of our put options with the equivalent amounts of cash and stock, we have

$$\begin{aligned}\Phi^{put,K+A}(S_T) - \Phi^{put,A}(S_T) &= \Phi^{call,K+A}(S_T) - \Phi^{call,A}(S_T) + (K+A-A)\Phi^{cash}(S_T) - (1-1)\Phi^{stock}(S_T)\\ &= \Phi^{call,K+A}(S_T) - \Phi^{call,A}(S_T) + K\Phi^{cash}(S_T).\end{aligned}$$

Hence, $\Phi(S_T)$ can be replicated with a portfolio containing one call option with strike price $K+A$, minus one call option with strike price K, and K units of cash.

(c) Corollary 14.3.3 refers to portfolios consisting only of stocks and cash. It does not apply to portfolios that are also allowed to contain derivatives.

15.5 We can write

$$\Phi^{bull}(S_T) = A + \max(S_T - A, 0) - \max(S_T - B, 0).$$

It may help to draw a picture, in the style of **15.4**. If we write $\Phi^{call,K}(S_T) = \max(S_T - K, 0)$ and recall that $\Phi^{cash}(S_T) = 1$ then we have

$$\Phi^{bull}(S_T) = A\Phi^{cash}(S_T) + \Phi^{call,A}(S_T) - \Phi^{call,B}(S_T).$$

So, in our (constant) hedging portfolio at time 0 we will need Ae^{-rT} in cash, one call option with strike price A, and minus one call option with strike price B.

15.6 The price computed for the contingent claim $\Phi(S_T) = S_t^\beta$ in **14.4** is

$$F(t, S_t) = S_t^\beta \exp\left(-r(T-t)(1-\beta) - \frac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right).$$

(Recall that Theorem 14.3.1 also tells us that there is a hedging strategy $h_t = (x_t, y_t)$ for the contingent claim with value $F(t, S_t)$.)

We have $F(t,s) = s^\beta \exp\left(-r(T-t)(1-\beta) - \frac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right)$ and we calculate

$$\begin{aligned}
\Delta_F &= \frac{\partial F}{\partial s}(t, S_t) = \beta S_t^{\beta-1} \exp\left(-r(T-t)(1-\beta) - \frac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right)\\
\Gamma_F &= \frac{\partial^2 F}{\partial s^2}(t, S_t) = \beta(\beta-1) S_t^{\beta-2} \exp\left(-r(T-t)(1-\beta) - \frac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right)\\
\Theta_F &= \frac{\partial F}{\partial t}(t, S_t) = S_t^\beta(1-\beta)(r + \tfrac{1}{2}\beta\sigma^2) \exp\left(-r(T-t)(1-\beta) - \frac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right)\\
\rho_F &= \frac{\partial F}{\partial r}(t, S_t) = -S_t^\beta(T-t)(1-\beta) \exp\left(-r(T-t)(1-\beta) - \frac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right)\\
\mathcal{V}_F &= \frac{\partial F}{\partial \sigma}(t, S_t) = -S_t^\beta\sigma\beta(T-t)(1-\beta) \exp\left(-r(T-t)(1-\beta) - \frac{1}{2}\sigma^2\beta(T-t)(1-\beta)\right).
\end{aligned}$$

15.7 (a) The underlying stock S_t has $\Delta_S = 1$ and $\Gamma_S = 0$. If we add -2 stock into our original portfolio with value F, then its new value is $V(t, S_t) = F(t, S_t) - 2S_t$, which satisfies $\Delta_V = \Delta_F - 2 = 0$.
The cost of adding -2 units of stock into the portfolio is $-2S_t$.

(b) After including an amount w_t of stock and an amount d_t of D, we have

$$V(t, S_t) = F(t, S_t) + w_t S_t + d_t D(t, S_t).$$

Hence, we require that

$$\begin{aligned}0 &= \Delta_F + w_t + d_t\Delta_D = 2 + w_t + d_t,\\ 0 &= \Gamma_F + d_t\Gamma_D = 3 + 2d_t.\end{aligned}$$

The solution is $d_t = \frac{-3}{2}$ and $w_t = \frac{-1}{2}$.
The cost of the extra stock and units of D that we have had to include is $-\frac{3}{2}D(t, S_t) - \frac{1}{2}S_t$.

15.8 (a) It doesn't work because adding in an amount z_t of Z in the second step will (typically i.e. if $\Delta_Z \neq 0$) destroy the delta neutrality that we gained from the first step.

(b) Since $Z(t,S_t)=S_t$ we have

$$V(t,S_t)=F(t,S_t)+w_tW(t,S_t)+z_tS_t,$$

so we require that

$$\begin{aligned}0&=\frac{\partial V}{\partial s}=\Delta_F+w_t\Delta_W+z_t\\ 0&=\frac{\partial^2 V}{\partial s^2}=\Gamma_F+w_t\Gamma_W.\end{aligned}$$

The solution is easily seen to be

$$\begin{aligned}w_t&=-\frac{\Gamma_F}{\Gamma_W}\\ z_t&=\frac{\Delta_W\Gamma_F}{\Gamma_W}-\Delta_F\end{aligned}$$

(c) This idea works. As we can see from part (b), if we use stock as our financial derivative $Z(t,S_t)=S_t$, then $\Gamma_Z=0$. Hence, adding in a suitable amount of stock in the second step can achieve delta neutrality without destroying the gamma neutrality obtained in the first step.

15.9 Omitted (good luck).

Chapter 17

17.1 The degrees of the nodes, in alphabetical order, are $(1,1),(0,2),(2,2),(2,0),(1,1),(1,2),(1,0)$. This gives a degree distribution

$$\mathbb{P}[D_G=(a,b)]=\begin{cases}\frac{1}{7} & \text{for } (a,b)\in\{(0,2),(2,2),(2,0),(1,2),(1,0)\},\\ \frac{2}{7} & \text{for } (a,b)=(1,1).\end{cases}$$

Sampling a uniformly random and moving along it means that the chance of ending up at a node A is proportional to $\deg_{\text{in}}(A)$. Since the graph has 8 edges, we obtain

$$\begin{aligned}\mathbb{P}[O=n]&=\begin{cases}\frac{1}{8}+\frac{2}{8} & \text{for } n=0 \text{ (nodes D and G)}\\ \frac{1}{8}+\frac{1}{8} & \text{for } n=1 \text{ (nodes A and E)}\\ 0+\frac{2}{8}+\frac{1}{8} & \text{for } n=2 \text{ (nodes B,C and F)}\end{cases}\\ &=\begin{cases}\frac{3}{8} & \text{for } n=0\\ \frac{1}{4} & \text{for } n=1\\ \frac{3}{8} & \text{for } n=2\end{cases}\end{aligned}$$

17.2 Node Y can fail only if the cascade of defaults includes $X\to D\to Y$. Given that X fails, the probability that X fails is $\eta_3=\frac{1}{3}$, because D has 3 in-edges. Similarly, given that D fails, the probability that Y also fails is $\frac{1}{2}$. Hence, the probability that Y fails, given that X fails, is $\frac{1}{6}$.

17.3 Node Y can fail if the cascade of defaults includes $X\to B\to D\to Y$ or $X\to B\to C\to D\to Y$. Given that X fails, node B is certain to fail as well, since B has only one in-edge. Given that B fails, C is certain to fail for the same reason. Therefore, the edges (B,D) and (C,D) are both certain to default. For each of these edges, independently, there is a chance $\frac{1}{3}$ that their own default causes D to fail. The probability that D fails is therefore

$$\frac{1}{3}+\left(1-\frac{1}{3}\right)\times\frac{1}{3}=\frac{5}{9}.$$

Here, we condition first on if the link BD causes D to fail (which it does with probability $\frac{1}{3}$) and then, if it doesn't (which has probability $1-\frac{1}{3}$) we ask if the link BC, which fails automatically and causes failure of CD, causes D to fail.

Given that D fails, Y is certain to fail. Hence, the probability that Y fails, given that X fails, is $\frac{5}{9}$.

17.4 For any node of the graph (except for the root node), if its single incoming loan defaults, then its own probability of default if α, independently of all else. Hence, each newly defaulted loan leads to two further defaulted loans with probability α, and leads to no further defaulted loans with probability $1-\alpha$. Hence, the defaulted loans form a Galton-Watson process (Z_n) with off-spring distribution G, given by $\mathbb{P}[G=2]=\alpha$ and $\mathbb{P}[G=0]=1-\alpha$, with initial state $Z_0=2$ (representing the two loans which initially default when V_0 defaults).

The total number of defaulted edges is given by

$$S = \sum_{n=0}^{\infty} Z_n.$$

Combining Lemmas 7.4.7, 7.4.6 and 7.4.8, we know that either:

- If $\mathbb{E}[G] > 1$ then there is positive probability that $Z_n \to \infty$ as $n \to \infty$; in this case for all large enough n we have $Z_n \geq 1$, and hence $S = \infty$.
- If $\mathbb{E}[G] \leq 1$ then, almost surely, for all large enough n we have $Z_n = 0$, which means that $S < \infty$.

We have $\mathbb{E}[G] = 2\alpha$. It is clear that the number of defaulted banks is infinite if and only if the number of defaulted loans in infinite, which has positive probability if $\mathbb{E}[G] > 1$. So, we conclude that there is positive probability of a catastrophic default if and only if $\alpha > \frac{1}{2}$.

17.5 (a) The probability that both B and C fail is $\frac{1}{2}\frac{1}{2} = \frac{1}{4}$. Given this event, the probability that both D and E fail is $\frac{1}{2}(1 - \frac{3}{4}\frac{3}{4}\frac{3}{4}) = \frac{37}{128}$; here the first term $\frac{1}{2}$ is the probability of D failing (via its link to B) and the second term is one minus the probability of E not failing despite all its inbound loans defaulting. Given that D and E both fail, the probability that F fails is $1 - \frac{2}{3}\frac{2}{3} = \frac{5}{9}$.

Hence, the probability that every node fails is

$$\frac{1}{4}\frac{37}{128}\frac{5}{9} = \frac{185}{4608}.$$

(b) Our strategy comes in three stages: first work out the probabilities of all the possible outcomes relating to B and C; secondly, do the same for D and E; finally, do the same for F. Each stage relies on the information obtained in the previous stage.

Stage 1: The probability that B fails and C does not is $\frac{1}{2}\frac{1}{2} = \frac{1}{4}$. This is also the probability that C fails and B does not. The probability that both B and C fail is also $\frac{1}{4}$.

Stage 2: Hence, the probability that E fails and D does not is

$$\frac{1}{4}\left(\frac{1}{4}\frac{1}{2}\right) + \frac{1}{4}\left(\frac{1}{4}\right) + \frac{1}{4}\left(\frac{1}{2}\right)\left(\frac{1}{4} + \frac{3}{4}\frac{1}{4}\right) = \frac{19}{128}$$

The three terms in the above correspond respectively to the three cases considered in the first paragraph. Similarly, the probability that D fails and E does not is

$$\frac{1}{4}\left(\frac{1}{2}\frac{3}{4}\frac{3}{4}\right) + \frac{1}{4}(0) + \frac{1}{4}\left(\frac{1}{2}\right)\left(\frac{3}{4}\frac{3}{4}\frac{3}{4}\right) = \frac{63}{512}$$

and the probability that both D and E fail is

$$\frac{1}{4}\left(\frac{1}{2}\right)\left(\frac{1}{4} + \frac{3}{4}\frac{1}{4}\right) + \frac{1}{4}(0) + \frac{1}{4}\left(\frac{1}{2}\right)\left(1 - \frac{3}{4}\frac{3}{4}\frac{3}{4}\right) = \frac{65}{512}$$

.

Stage 3: Finally, considering these three cases in turn, the probability that F fails is

$$\frac{19}{128}\left(\frac{1}{3}\right) + \frac{63}{512}\left(\frac{1}{3}\right) + \frac{65}{512}\left(1 - \frac{2}{3}\frac{2}{3}\right) = \frac{371}{2304}.$$

Appendix D

Advice for revision/exams

There are two different exam papers, one for MAS352 and one for MAS452/6052. For both exams the rubric reads

> *Candidates should attempt ALL questions. The maximum marks for the various parts of the questions are indicated. The paper will be marked out of 100.*

Within these notes, material marked with a (Δ) is examinable only for MAS452/6052, and is non-examinable for MAS352. Material marked with a ($\star$) is non-examinable for everyone.

- You will be asked to solve problems based on the material in these notes. There will be a broad range of difficulty amongst the questions. Some will be variations of questions in the assignments/notes, others will also try to test your ingenuity.
- You may be asked to state important definitions and results (e.g. more than one past exam has asked for definition of Brownian Motion).
- You will not be expected to reproduce long proofs from memory. You are expected to have followed the techniques within the proofs (e.g. Ito's formula, conditional expectation rules) when they are present, and to be able to use these techniques in your own problem solving.
- There are marks for attempting a suitable method, and for justifying rigorous mathematical deductions, as well as for reaching a correct conclusion.
- If you apply an important result that has a name e.g. 'the Martingale Convergence Theorem' you should mention that name, or something similar e.g. 'by mart. conv.' or 'by the MCT'.
- Practice using the formula sheet to help you solve questions. It contains lots of useful formulae!

Revision activities

The most important activities:

1. Check and mark your solutions to assignment questions.
2. Learn the key definitions, results, and examples.
3. Do the past exam papers, and mark your own solutions.

Other very helpful activities:

4. Work through, and check your solutions, to non-challenge questions in the notes.

Of course, you should have been working on these questions throughout the year, which is why they are lower priority now. You do *not* need to look at the challenge questions as part of your revision – these are intended only to offer a serious, time consuming challenge to strong students.

In all cases, you are welcome to come and discuss any questions/comments/typos. Please email to arrange a convenient time.

Appendix E

Formula sheet (part two)

The formula sheet displayed on the following two pages will be provided in the exam.

MAS352/452/6052 – Formula Sheet – Part Two

Where not explicitly specified, the notation used matches that within the typed lecture notes.

The normal distribution

$Z \sim N(\mu, \sigma^2)$ has probability density function $f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$.

Moments: $\mathbb{E}[Z] = \mu$, $\mathbb{E}[Z^2] = \sigma^2 + \mu^2$, $\mathbb{E}[e^Z] = e^{\mu + \frac{1}{2}\sigma^2}$.

Ito's formula

For an Ito process X_t with stochastic differential $dX_t = F_t\,dt + G_t\,dB_t$, and a suitably differentiable function $f(t,x)$, it holds that

$$dZ_t = \left\{ \frac{\partial f}{\partial t}(t, X_t) + F_t \frac{\partial f}{\partial x}(t, X_t) + \frac{1}{2} G_t^2 \frac{\partial^2 f}{\partial x^2}(t, X_t) \right\} dt + G_t \frac{\partial f}{\partial x}(t, X_t)\,dB_t$$

where $Z_t = f(t, X_t)$.

Geometric Brownian motion

For deterministic constants $\alpha, \sigma \in \mathbb{R}$, and $u \in [t, T]$ the solution to the stochastic differential equation $dX_u = \alpha X_u\,dt + \sigma X_u\,dB_u$ satisfies

$$X_T = X_t e^{(\alpha - \frac{1}{2}\sigma^2)(T-t) + \sigma(B_T - B_t)}.$$

The Feynman-Kac formula

Suppose that $F(t,x)$, for $t \in [0,T]$ and $x \in \mathbb{R}$, satisfies

$$\begin{aligned} \frac{\partial F}{\partial t}(t,x) + \alpha(t,x)\frac{\partial F}{\partial x}(t,x) + \frac{1}{2}\beta(t,x)^2 \frac{\partial^2 F}{\partial x^2}(t,x) - rF(t,x) &= 0 \\ F(T,x) &= \Phi(x). \end{aligned}$$

If X_u satisfies $dX_u = \alpha(u, X_u)\,dt + \beta(u, X_u)\,dB_u$, then

$$F(t,x) = e^{-r(T-t)} \mathbb{E}_{t,x}\left[\Phi(X_T)\right].$$

The Black-Scholes model

The Black-Scholes model is parametrized by the deterministic constants r (continuous interest rate), μ (stock price drift) and σ (stock price volatility).

The value of a unit of cash C_t satisfies $dC_t = rC_t\,dt$, with initial value $C_0 = 1$.

The value of a unit of stock S_t satisfies $dS_t = \mu S_t\,dt + \sigma S_t\,dB_t$, with initial value S_0.

At time $t \in [0,T]$, the price $F(t,S_t)$ of a contingent claim $\Phi(S_T)$ (satisfying $\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)] < \infty$) with exercise date $T > 0$ satisfies the Black-Scholes PDE:

$$\begin{aligned}\frac{\partial F}{\partial t}(t,s) + rs\frac{\partial F}{\partial s}(t,s) + \frac{1}{2}s^2\sigma^2\frac{\partial^2 F}{\partial s^2}(t,s) - rF(t,s) = 0,\\ F(T,s) = \Phi(s).\end{aligned}$$

The unique solution F satisfies

$$F(t,S_t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\Phi(S_T)\,|\,\mathcal{F}_t]$$

for all $t \in [0,T]$. Here, the 'risk-neutral world' $\mathbb{Q}$ is the probability measure under which S_t satisfies

$$dS_t = rS_t\,dt + \sigma S_t\,dB_t.$$

The Gai-Kapadia model of debt contagion (MAS452/6052 only)

A financial network consists of banks and loans, represented respectively as the vertices V and (directed) edges E of a graph G. An edge from vertex X to vertex Y represents a loan owed by bank X to bank Y.

Each loan has two possible states: healthy, or defaulted. Each bank has two possible states: healthy, or failed. Initially, all banks are assumed to be healthy, and all loans between all banks are assumed to be healthy.

Given a sequence of contagion probabilities $\eta_j \in [0,1]$, we define a model of debt contagion by assuming that:

(†) For any bank X, with in-degree j if, at any point, X is healthy and one of the loans owed to X becomes defaulted, then with probability η_j the bank X fails, independently of all else. All loans owed by bank X then become defaulted.

Starting from some set of newly defaulted loans, the assumption (†) is applied iteratively until no more loans default.